

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:45:35 ; Search time 4020 Seconds
(without alignments)
8425.422 Million cell updates/sec

Title: US-10-049-710A-1
Perfect score: 699
Sequence: 1 atggcagtcgcacagacg.....aggcgcctccgacagtag 699

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	699	100.0	699	6	BD105336
2	699	100.0	1755	1	STMBARA
3	267.2	38.2	2234	1	AB001683 Streptomy
4	229	32.8	6727	1	SCO007731
5	229	32.8	290850	1	SCO939127
6	213.6	30.6	1328	1	STMAFR
7	211.4	30.2	33064	2	AY338477
8	211.4	30.2	210614	1	AB088224
9	199.2	28.5	1601	1	AB021882
10	197.8	28.3	654	1	AB121071
11	197.8	28.3	6997	1	AB126048 Kitasatos
12	194.4	28.0	901	1	AY256849
13	194.4	27.8	2637	1	AF156161
14	180.8	25.9	92294	1	AY117439 Streptomy
15	173.4	24.8	1101	1	AB106894
16	163	23.3	1129	1	AY026762
17	157.8	22.6	104578	1	AY502076 Rhodococc
18	157.4	22.5	300800	1	AP005036
19	154.8	22.1	10467	1	AF145049

20	145.8	20.9	1340	1	AB000385	AB000385 Streptomy
21	145.6	20.8	295150	1	SC0939126	AL939126 Streptomy
22	138	19.7	799	1	AB000384	AB000384 Streptomy
23	126	18.0	12905	1	AF055922	AF055922 Streptomy
24	112.4	16.1	300800	1	AP005036	AP005036 Streptomy
25	93.4	13.4	300550	1	AP005030	AP005030 Streptomy
26	92.2	13.2	104578	1	AY502076	AY502076 Rhodococc
27	90.6	13.0	1498	1	AB001609	AB001609 Streptomy
28	88.6	12.7	1819	6	AX708269	AX708269 Sequence
29	88.6	12.7	1819	6	AX708269	AX708269 Sequence
30	85.8	12.3	9521	6	AX188771	AX188771 Sequence
31	85.8	12.3	9521	6	AX188772	AX188772 Sequence
32	85.8	12.3	22818	1	SC0276673	AJ276673 Streptomy
33	85.8	12.3	178073	1	SC0590464	AL590464 Streptomy
34	85.6	12.2	110000	1	AP006618_50	Continuation (51 o
35	82	11.7	3729	1	AB035547	AB035547 Streptomy
36	79.8	11.4	321250	1	SC0939111	AL939111 Streptomy
37	78.8	11.3	110000	1	AP006618_06	Continuation (7 of
38	78.8	11.3	110000	1	AP006840_24	Continuation (25 o
39	78.2	11.2	4257	6	AR144708	AR144708 Sequence
40	78.2	11.2	4257	6	AR145616	AR145616 Sequence
41	78.2	11.2	4257	6	AR526866	AR526866 Sequence
42	78.2	11.2	6633	14	HEHSVIG3	X06461 Herpes simp
43	78.2	11.2	26245	14	HS1US	L00036 Human herpe
44	78.2	11.2	26245	14	HS1US	L00036 Human herpe
45	78.2	11.2	152261	14	HE1CG	X14112 Human herpe

ALIGNMENTS

RESULT 1
BD105336
LOCUS
DEFINITION Method for induction of gene expression in plants and plants
thereof
ACCESSION BD105336
VERSION BD105336.1 GI:22650910
KEYWORDS WO 0196581-A/1.
SOURCE Streptomyces virginiae
ORGANISM Streptomyces virginiae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomycetes.
REFERENCE 1 (bases 1 to 699)
AUTHORS Shimmyo,A., Kato,K., Yamada,Y., Nihira,T. and Shindo,T.
TITLE Method for induction of gene expression in plants and plants
thereof
JOURNAL Patent: WO 0196581-A 1 20-DEC-2001;
KANAKA CORP,ATSUHIKO SHIMMYO,KO KATO,YASUHIRO YAMADA,TAKUYA NIHIRA,
TAKUYA SHINDO
COMMENT OS Streptomyces virginiae
PN WO 0196581-A/1
PD 20-DEC-2001
PF 15-JUN-2001 WO 2001JP005096
PR 15-JUN-2000 JP 00P 190466
PI ATSUHIKO SHIMMYO,KO KATO,YASUHIRO YAMADA,TAKUYA NIHIRA,TAKUYA
SHINDO
PC C12N15/82,A01H5/00,C12N5/10
CC Method for induction of gene expression in plants and plants
thereof
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..699
/organism="Streptomyces virginiae"
/mol_type="genomic DNA"
/db_xref="taxon:1961"

FEATURES
source

ORIGIN

Query Match 100.0%; Score 699; DB 6; Length 699;
Best Local Similarity 100.0%; Pred. No. 5.9e-75;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCAGTGGCGACACGAAACGGGTGGGAGTGGCGACAGAAACGGGCCCTTCGCGACGGCGGAG 60
Db 1 ATGGCAGTGGCGACACGAAACGGGTGGGAGTGGCGACAGAAACGGGCCCTTCGCGACGGCGGAG 60
QY 61 GCGATCGTGGGCGACGGCGCTCGGTCTTCGACAGTACGGGTTCGAGGCGGCGACAGTG 120
Db 61 GCGATCGTGGGCGACGGCGCTCGGTCTTCGACAGTACGGGTTCGAGGCGGCGACAGTG 120
QY 121 GCAGAGATCCTCTTCGCGGGGCTCGGTCAACAGGGCGGATGATGTACTTCCACTTCGCTTCC 180
Db 121 GCAGAGATCCTCTTCGCGGGGCTCGGTCAACAGGGCGGATGATGTACTTCCACTTCGCTTCC 180
QY 181 AAGGAAGAGTGGCCCGCGCGGTGCTGGCGGAGCAGACCTGCACTGGCGGTGCGCGAA 240
Db 181 AAGGAAGAGTGGCCCGCGCGGTGCTGGCGGAGCAGACCTGCACTGGCGGTGCGCGAA 240
QY 241 TCCGGCTCCAAGCGCGAGGACTGTGTAGACCTTACCATGCTGTGTCGCCACGGCATGCTG 300
Db 241 TCCGGCTCCAAGCGCGAGGACTGTGTAGACCTTACCATGCTGTGTCGCCACGGCATGCTG 300
QY 301 CACGATCCGATCTCGCGGGCGGCGACCGCGCTCGCACTGGACAGGGGGCGGTGGACTTC 360
Db 301 CACGATCCGATCTCGCGGGCGGCGACCGCGCTCGCACTGGACAGGGGGCGGTGGACTTC 360
QY 361 TCGACGCCAACCCGTTCCGCGAGTGGGCGACATCTGCGCCAGTCTCTGCGCGAGGCA 420
Db 361 TCGACGCCAACCCGTTCCGCGAGTGGGCGACATCTGCGCCAGTCTCTGCGCGAGGCA 420
QY 421 CAGGAACGGGGGAGGTCTTCGACACGTGMAACCCGAAAGACCGGCGACTTCATCGTC 480
Db 421 CAGGAACGGGGGAGGTCTTCGACACGTGMAACCCGAAAGACCGGCGACTTCATCGTC 480
QY 481 GGTGCTTCCACGGGCTCCAGGCGGTCTCCGCGGTCACTCCGACCGCGAGGACCTTCGCG 540
Db 481 GGTGCTTCCACGGGCTCCAGGCGGTCTCCGCGGTCACTCCGACCGCGAGGACCTTCGCG 540
QY 541 CACCGATCTCGGTATGTGAACACAGTGTGTCGCCAGCATGTCGCGCGGCTTCATGCTG 600
Db 541 CACCGATCTCGGTATGTGAACACAGTGTGTCGCCAGCATGTCGCGCGGCTTCATGCTG 600
QY 601 ACTTGATCGAAACCGGCGAGGAGCGGATCGGAGGTCTCGCGCGCGCGCGCGCGCC 660
Db 601 ACTTGATCGAAACCGGCGAGGAGCGGATCGGAGGTCTCGCGCGCGCGCGCGCGCC 660
QY 661 GAGGCTGGGAGGCTTCGAGGCGGCTTCGAGCGGCTTCGAGCGAGTAG 699
Db 661 GAGGCTGGGAGGCTTCGAGGCGGCTTCGAGCGGCTTCGAGCGAGTAG 699
```

```
RESULT 2
STMBARA 1755 bp DNA linear BCT 07-FEB-1999
LOCUS Streptomyces virginiae barA gene for virginiae butanolide receptor,
DEFINITION complete cds.
ACCESSION D32251.1 GI:961502
VERSION Streptomyces virginiae butanolide receptor; barA.
KEYWORDS Streptomyces virginiae
SOURCE Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
ORGANISM Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1. (bases 1 to 1755)
AUTHORS Okamoto, S., Nakamura, K., Nihira, T. and Yamada, Y.
TITLE Evidence that VbraA is not the virginiae butanolide binding protein
and reidentification of the true binding protein
J. Biol. Chem. 270 (20), 12319-12326 (1995)
JOURNAL 95263588
MEDLINE 7744885
PUBMED 2 (bases 1 to 1755)
REFERENCE Nihira, T.
AUTHORS Direct Submission
TITLE Submitted (19-JUL-1994) Takuya Nihira, Osaka University, Department
```

```
COMMENT
of Biotechnology; 2-1 Yamadaoka, Suita, Osaka 565, Japan
(Tel:06-879-7433, Fax:06-879-7448)
Submitted (19-Jul-1994) to DDBJ by:
Takuya Nihira
Osaka University
Department of Biotechnology
2-1 Yamadaoka
Suita, Osaka 565
Japan
Phone: 06-879-7433
Fax: 06-879-7448.
FEATURES
source
1. 1755
/organism="Streptomyces virginiae"
/mol_type="genomic DNA"
/isolate="WAPP 10-06014"
/db_xref="taxon:1961"
442.445
/citation=[1]
/evidence=not_experimental
454.1152
/gene="barA"
454.1152
/gene="barA"
/function="binding of virginiae butanolides"
/standard_name="BarA"
/citation=[1]
/codon_start=1
/evidence=experimental
/transl_table=1
/product="virginiae butanolide receptor"
/protein_id="BAA06981.1"
/db_xref="GI:961503"
/translation="MAVHERVAVRQRAVTRQAI VRAAASVFDEYFEATVAEIL
SRASVTKGAMYFHPASKEELARGLAEQTLHVAVPESCKAQELVDLMLVAHGLHD
FILRAGTKALDQGVDFSDPFGEWDICQALLAEQERGEVLPVHNPKTKGDFIV
GCTTQLQVSRVTSRDRLGHRISVMNHNVLPSIVPASMLTWIETGEERIGKVAANA
AAEAASEAASDE"
ORIGIN
Query Match 100.0%; Score 699; DB 1; Length 1755;
Best Local Similarity 100.0%; Pred. No. 4.6e-75;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCAGTGGCGACACGAAACGGGTGGGAGTGGCGACAGAAACGGGCCCTTCGCGACGGCGGAG 60
Db 454 ATGGCAGTGGCGACACGAAACGGGTGGGAGTGGCGACAGAAACGGGCCCTTCGCGACGGCGGAG 513
QY 61 GCGATCGTGGGCGACGGCGCTCGGTCTTCGACAGTACGGGTTCGAGGCGGCGACAGTG 120
Db 514 GCGATCGTGGGCGACGGCGCTCGGTCTTCGACAGTACGGGTTCGAGGCGGCGACAGTG 573
QY 121 GCAGAGATCCTCTTCGCGGGGCTCGGTCAACAGGGCGGATGATGTACTTCCACTTCGCTTCC 180
Db 574 GCAGAGATCCTCTTCGCGGGGCTCGGTCAACAGGGCGGATGATGTACTTCCACTTCGCTTCC 633
QY 181 AAGGAAGAGTGGCCCGCGCGGTGCTGGCGGAGCAGACCTGCACTGGCGGTGCGCGAA 240
Db 634 AAGGAAGAGTGGCCCGCGCGGTGCTGGCGGAGCAGACCTGCACTGGCGGTGCGCGAA 693
QY 241 TCCGGCTCCAAGCGCGAGGACTGTGTAGACCTTACCATGCTGTGTCGCCACGGCATGCTG 300
Db 694 TCCGGCTCCAAGCGCGAGGACTGTGTAGACCTTACCATGCTGTGTCGCCACGGCATGCTG 753
QY 301 CACGATCCGATCTCGCGGGCGGCGACCGGCTTCGACAGTGGACACGGGGCGGTGGACTTC 360
Db 754 CACGATCCGATCTCGCGGGCGGCGACCGGCTTCGACAGTGGACACGGGGCGGTGGACTTC 813
QY 361 TCGACGCCAACCCGTTCCGCGAGTGGGCGACATCTGCGCCAGTCTCTGCGCGAGGCA 420
Db 814 TCGACGCCAACCCGTTCCGCGAGTGGGCGGAGCATCTGCGCGGAGGCA 873
QY 421 CAGGAACGGGGGAGGTCTTCGCGAGTGGGAGGTCTCGCGCGCGCGCGCGCGCC 480
```

```
Db 874 CAGGAACGGGGGAGGTCTTCGACGTAACCCGAAAGACCGGCGACTTCATCGTC 933
Qy 481 GGCTGCTTACCCGGGCTCAGAGCGGTCTCCCGGGTCACTCCGACCGCCAGGACCTCGGC 540
Db 934 GGCTGCTTACCCGGGCTCAGAGCGGTCTCCCGGGTCACTCCGACCGCCAGGACCTCGGC 993
Qy 541 CACCGATCTCGTGTATGTGAACACGCTCTGCCAGATCTGCGCGGCTCCATGCTG 600
Db 994 CACCGATCTCGTGTATGTGAACACGCTCTGCCAGATCTGCGCGGCTCCATGCTG 1053
Qy 601 ACCTGATCGAACCAGGAGCGGATCGGAAGTCTCGCGCGCGGCGGCGGCGGCG 660
Db 1054 ACCTGATCGAACCAGGAGCGGATCGGAAGTCTCGCGCGCGGCGGCGGCGGCG 1113
Qy 661 GAGGCTCGGAGGCTCCGAGGCGGCTCCGACGAGTAG 699
Db 1114 GAGGCTCGGAGGCTCCGAGGCGGCTCCGACGAGTAG 1152

RESULT 3
AB001683 2234 bp DNA linear BCT 22-AUG-1997
LOCUS Streptomyces sp. gene for FarX, FarA, complete cds.
DEFINITION Streptomyces sp.
ACCESSION AB001683
VERSION AB001683.1 GI:2342428
KEYWORDS FarX, FarA.
SOURCE Streptomyces sp.
ORGANISM Streptomyces sp.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE
AUTHORS Waki,M., Nihira,T. and Yamada,Y.
TITLE Cloning and characterization of the gene (farX) encoding the
receptor for an extracellular regulatory factor (IM-2) from
Streptomyces sp. strain FRI-5
JOURNAL J. Bacteriol. 179 (16), 5131-5137 (1997)
MEDLINE 97405912
PUBMED 9260956
REFERENCE 2 (bases 1 to 2234)
AUTHORS Waki,M.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-1997) Miyoko Waki, Osaka University, Graduate
School of Engineering, Department of Biotechnology; Yamadaoka 2-1,
Suita, Osaka 565, Japan (E-mail:waki.yam@stu.bio.eng.osaka-u.ac.jp,
Tel.:+81-6-879-7433, Fax:+81-6-879-7432)
FEATURES
source
1..2234
/organism="Streptomyces sp."
/mol_type="genomic DNA"
/strain="FRI-5"
/db_xref="taxon:1931"
440..1315
/gene="farX"
440..1315
/gene="farX"
/codon_start=1
/transl_except="(pos:440..442,aa:Met)
/transl_table=11
/product="FarX"
/protein_id="BAA21858.1"
/db_xref="GI:2342429"
/translations="MHRTSTAQVLLTMDQLDARFSTVTRWPLSHAFPTVPVGDGY
DPMCAETRIQAVLLGHAEFVFGHOFVLDLSVVRPELLRVLGLVPATVDLAI
CWEIKRAGLSGLGEAVVRDGOVATGASVCTSPAVQRIPEHVLTPHRLPL
PLTAPAPQSVARLSPTDVLSPDLRENWQLRVDNHPVLFHWDHPVGMVLMMA
RQAASALGRPSFMPPLGAGEFKRYVELDAPCIESERLFDQVFGAEVVRVTGHQ
ELTFVGTVTASSGY"
1500..2165
/gene="farA"
1500..2165
/gene="farA"
/codon_start=1
```

```
/transl_except="(pos:1500..1502,aa:Met)
/transl_table=11
/product="FarA"
/protein_id="BAA21859.1"
/db_xref="GI:2342430"
/translations="NAEQVRAIRTRQAILSAARVFDERGYOAAITSEILTAVGVTKG
ALYFHFQSKEDLAQVILTAQNEIDLILPERPAKQEVVDVAMLHTRLFNPMVRAGVR
LSLDYNAGGLDRSAPFRNWDKFTDLLEKAQAGELLPHVPAETADVITGAYGVQS
MSQALTEHQDLGQVRNALLHLMPSIAOPSVLASLHLSRAEVEVYLEARQAREAD
EBD"

ORIGIN
Query Match 38.2%; Score 267.2; DB 1; Length 2234;
Best Local Similarity 63.0%; Pred. No. 2.6e-23;
Matches 413; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

Qy 33 ACAGGAACGGGGCTCCGACCGGAGCGGATCGTGGGGGAGCGGCTCGGTCTTCGA 92
Db 1508 ACAGGTCGAGCATCCGACCGGCGGAGTCTTGAGCGCGGCGGAGGTGTTTGA 1567
Qy 93 CGAGTACGGGTTGAGGCGGCGCACAGTGGCAGAGATCTCTTCGGGGGCTTCGTACCAA 152
Db 1568 CGAACGGGGCTACCGAGCGGCGCACCATCTCGGAGATCTCTCACGTCGCGGGGTGACCAA 1627
Qy 153 GGGCGGATGTACTTCCACTTCGCTTCCAAGGAGAGCTGGCCGCGGCGGTGCTGCCGA 212
Db 1628 GGGCGGCTGTACTTCCACTTCGAGTCCAAGGAGAGCTGGCCGCGGCGGTGCTGCCGC 1687
Qy 213 GCAGACCTTCGACGTCGGGCTGCCGAATCCGCTCCAAGGCGCAGGAACTGGTAGACCT 272
Db 1688 GCAGAACGAGGATCTTCTCTCTGAGGCGCCCGCAAACTCCAGAGGTGGTGTGATGC 1747
Qy 273 CACCATGCTGGTGGCCACCGGATGTGCACGATCCGATCTCTGGGGCGGCGACGGGCT 332
Db 1748 CGTCATGCTGCACACCCACCGCTGGCGACCAACCCATGTTCCGCGCGGCGTCCGGCT 1807
Qy 333 CGCACTGGACGAGGGGCGGTGGACTTCTCCAGCGCAACCCGTTGGCGAGTGGGGCGA 392
Db 1808 GAGCCTCGACGTAACCGCGGGGGGCTCGACCGCAGCGTTCGTTCCGCAACTGGGTCGA 1867
Qy 393 CATCTCGGCGGAGTCTCTGGCGGAGGCACAGGAACGGGGGAGGTGCTTCCGACACGTAA 452
Db 1868 CAGTTTCACGACCTCTCTGGAGAGGCGCCAGGCCAGGCGGAACTGTCGCCACGTTGT 1927
Qy 453 CCGGAAAAGACGGCGACTTCATCTCGGTGCTTTCACGGGCTTCAGGCGGTCTCCCG 512
Db 1928 ACCGGCGGAGACGGCGGCGTTCATCACCGGTCTACGGCGGCGTCCAGTCCATGTCCCA 1987
Qy 513 GGTCACTCGGACCGCGAGACTTCGGCCACCGGATCTCGGTGATGTGGAACACCGTCT 572
Db 1988 GGCGCTCACGAGCACCGGACCTCGGGCAGCGGGTCAACGCGCTGTCGCGCACCTCAT 2047
Qy 573 GCCCAGCATCTGTCGGCGGCTCCATGCTGACCTGGATCGAAACCGCGGAGAGCGGATCGG 632
Db 2048 GCGGACATCGCCAGCCCTCGGTCTCTCGCTCCCTGACCTCGGCGAGAGCGGGCGGA 2107
Qy 633 GAAGTTCGCGGGCGGCGGCGGAGCCCGAGGCTCGGAGGCTTCGAGGCGCGCT 688
Db 2108 AGAGGTCTACCTCGAAGCCCGGCGAGTGTGCGCGGAGCAGCGGAGCGGAGAGACT 2163

RESULT 4
SC0007731 6727 bp DNA linear BCT 24-AUG-1998
LOCUS Streptomyces coelicolor scbr gene, scbA gene, ORFs A,B,X & Z.
DEFINITION Streptomyces coelicolor scbr gene, scbA gene; scbr gene.
ACCESSION AJ007731
VERSION AJ007731.1 GI:3425857
KEYWORDS gamma-butyrolactone binding protein; scbA gene; scbr gene.
SOURCE Streptomyces coelicolor A3(2)
ORGANISM Streptomyces coelicolor A3(2)
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1
```

AUTHORS Takano, E., Chakraborty, R., Nihira, T., Yamada, Y. and Bibb, M.
TITLE Characterisation of *scbR* and *scbA* genes involved in gamma-butyrolactone binding and synthesis in *Streptomyces coelicolor*

JOURNAL Submitted (11-AUG-1998) Takano E., Dept. Genetics, John Innes Institute, Colney lane, Norwich, NR4 7UH, UK

FEATURES Location/Qualifiers
1..6727

gene

CDs

```

/codon_start=1
/transal_table=11
/product="histidine kinase"
/protein_id="CAA07625.1"
/db_xref="GI:3425858"
/db_xref="GOA:O86849"
/db_xref="UniProt/TREMBL:O86849"
/transalation="WGTPTSPENRRGRTTLCSDLPSPSRRTVRHFPCPTPAKRD
LAGSHRGGRFTTLDGSLRSFGRAGOTLTGHPGLPADGRDPAAGVPTDGRDGD
HWAEPRRLRLAEVHDELGTALSTAHHIELHAETGCTPHLDAARHSLREMAVARRL
TGEQAOCTVPLPVAQVGEFAATTPPTGTEVKRITGDERLLSDVCRRELFPLAVREAL
HNAFHAAHDARVTVTLRPRWHAAGIVDDGVGDFDAVLPFGHRAVPGFLPSMTDRIED
VGWRLIIVSPAGGCTHIDVHLPRKVTAPRT"
complement(1000...1854)

```

gene

CDs

```

/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA07626.1"
/db_xref="GI:3425859"
/db_xref="UniProt/TREMBL:O86850"
/translation="MLVTGVTVGVGVROVAILHARTGPVRVILARRPVRVTVTPNVVE
CAGEYADRGLGRALNGVAAFLVNTSATEPDRRPAARAAEAGRHVLKLSLVANGE
GACGDDPTVTRGENSRVAPESGLDWTFRATSTWNTLSWAPAIRSDGVVVALYKGYSP
VACQDPRTIAEYVALRTPEESGRCAYALSPEELTARTQARLSVEVLGRSLAPFELG
LEAARQMLKRPVPAVDFLESARQAGAAQVVFIVRETVTGPVPPRPIWADHA
EAFRQGG"
/complement(1970..2914)
/gene="scbA"
/complement(1970..2914)
/gene="scbA"
/function="involved in gamma-butyrolactone synthesis"
/codon_start=1
/transl_table=11

```

```

/od_xref="GI:3425860"
/db_xref="UniProt:O66851"
/translation="MPEAVLINSASDANSTEQTALPQMALVHRTVQDAFPVSMIP
KGGDRSVTAIVLPHDPFFPAVGHGRDHPDLIAETLQKQGVQVQVAVRAGLAATGTA
ATLUDYTHLHLDHVGSGEVALEVEVACSKFQPGQVQVQVQVQVQVQVQVQVQVQVQV
TTPQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV
DTSHPTLQFQRNDHVGMLLEAARQAACTVGPAPFVPSIGGTRFVRYAEFDSPCWI
QATVRGPAAGITTVRVTHQDGSUVFLTILSGPAFSG"
3032..3679
/genes="scbr"
3032..3679
/genes="scbr"
/codon_start=1
/evidence=experimental
/transl_table=11

```

gene

CDS

```
/codon_start=1
/evidence=experimental
/transl_table=11
```

```

/product="gamma-butyrolactone binding protein"
/protein_id="CAA07628.1"
/db_xref="GI:3425861"
/db_xref="UniProt:O86852"
/db_xref="UniProt/T:EMBL:O86852"
/tranlation="MAKQDRAISTROTILDAAQVFEKQGYQAATITELIKLVAGTVKGR
ALYFPHQSKELALGVFDAAQBPQAPQPEQLRIQLIDMGLFCHRLRTNVAARVGR
LSMDQAHGDRGRGPFDDAQETPQAPQPEQLRIQLIDMGLFCHRLRTNVAARVGR
VSQTVSQDLEHRYALQLHILPAITAVPSVLAALDLSERGARLAAELAAPTGKD"
complement(3795..4529)
/gene="orfX"
/codon_start=1
/gene="orfX"
/product="3-ketoacyl-ACP/CoA reductase"
/protein_id="CAA07629.1"
/db_xref="GI:3425862"
/db_xref="GOA:O86853"
/db_xref="UniProt/T:EMBL:O86853"
/tranlation="MTGSGRGIGRTALRLAADGALVAVHYGSSAAAEVETVIBSS
GGQALATRAELGVGDAAALYAAFDAGMEFGVPPFEDILVNNAGVSGSGRTEVTEEE
VFDRLVAVNRAPIFVQHGKELRDGGRIINSSAATRAFPESIGYATMGKAVDIL
TLAARQLGERTGVNAVAFGVTEKRLDGGRIINSSAATRAFPESIGYATMGKAVDIL
VWAFASDSSRWITGQYVDATGGTIL"
complement(4746..6446)
/gene="orfZ"
/codon_start=1
/gene="orfZ"
/product="hypothetical protein"
/protein_id="CAA07630.1"
/db_xref="GI:3451475"
/db_xref="UniProt/T:EMBL:O86854"
/tranlation="MWRVRCGRILVGRRRRRPKASRPGGAAPAEWSRSPAAQRQPA
GGFTADARAVRDEFAAGAFRLPPENDQAEPLPAGEQGTSSAGGTAHAERAPAS
PGTAHAPSPGADTQGSVSVVRGRTVSAPPKAKARFMSMAPDRDGDRLRTFFAL
PPREAAGEPAAVTWLNAGWNTALEGALDAAERKGRVGAEHGHVDATVTPGLVLVAY
VGRSGSRPYRVQRLDGLTDSMDRFLDAAVERPGHIAALDGLFHSIADLRGVP
LIFQGLDAPRCSPDGSGHPCKRHAALCYQTARLLDAPFVLLLRGGERALLDALS
RRNAARARAAQDRGPGPLPGVRAGAAITPRALPLPAPLPAPHPHQEPYAPAAPGG
PYPPPLDHLATAAARAHAALLTGRDPVGLGLTQDAVRLAAARPGSGLTAGTRALYA
SLASACRGDTAEALARVAAWRQGLAGLDVLEEPDPAPGRFDRARMMLAADLPAPR
PWRNLTTHPRGHVQLRGTRGLWYAYESELGRDWMFRGTPTDLDLVGALTGLGGPDP

```

ORIGIN

	Query Match	32.8%; Score 229; DB 1; Length 6727;
	Best Local Similarity	60.8%; Pred. No. 7.3e-19;
Matches	373; Conservative	0; Mismatches 240; Indels 0; Gaps 0;
Qy	34	CAGAAAGGGCGCTCCGACGGGCAGGGCATCGTGGGGCAGCCGCTCGGTCTTCGAC 93
Dd	3041	CAGAACGGGCGATCCGCAACGGGCAGACATCCTGGACGCCCGCGCAGGTCTTCGAG 3100
Qy	94	GAGTAGGGTTCCAGAGCGCCACAGTGGCAGAGATCCTCTCGGGGCGCTCGGTCAACAAG 153
Dd	3101	AAGCAGGGCTACCAGCTGCCCCAATACGGAGATCCTCAAGTGCGGGGTGACCAAG 3160
Qy	154	GGGCGCATGTACTTCCACTTCGCTTCCAAGGAAGAGCTGGCCCGCGCGCTGTGGCGCGAG 213
Dd	3161	GGAGCCCTCTACTTCCACTTCCAGTCCAGGAAGAAGACTGGCGTGGGGCTCTTCGACGCC 3220
Qy	214	CAGACCGTGCACGTGGCGGTGCGGAATCCGGTCCAAAGCGCAGGAATGTGTAGACCTC 273
Dd	3221	CAGGAACCAACACAGGCGGTTCGGAGCAACCCTCCGGTGTCAAGAACTCATCGACATG 3280
Qy	274	ACCATGCTGTCGCCCAACGGCATGCTGCAGATCCGATCCTGCGGGGGGGCACGCGGCTC 333
Dd	3281	GGCATGTGTGTCTGTCA CGCTTTCGCA CAAACGTGTGCGCCCGGCGGTGCGGCTC 3340
Qy	334	GCACCTGGACAGGGGGCGGTGGACTTCTCCAGCGCCAAACCGTTCGGCGAGTGGGGCGAC 393

Db 3341 TCATGACACAGAGCGGCACGGTCTCGATCGCCGAGGACCCCTTCGTCGTGGCACGAG 3400
Qy 394 ATCTGCGCCAGCTCTCTGGCGGAGGACACGAGAACGGGGGAGGTGCTTCGGCAGCTGAAC 453
Db 3401 ACACCTCTGAAGCTGCTGAACACGAGGCCAAGGAGACGGTGAGTGTCTGCCCCATGTGTC 3460
Qy 454 CCGAAAAAGACCGGACATTCATCTGCTGGCTGCTTCACCGGGTCCAGCGGTCTCCCGG 513
Db 3461 ACCACGATCGGCCGATCTCTACGTGGGCACGCTTCGCCGGGATACAGGTCTGTGTC 3520
Qy 514 GTCACCTCCGACCGCCAGACCTCGCCACCGGATCTCGTGATGGAACACGAGTGTG 573
Db 3521 ACGGTGAGGACTACAGGACCTCGAACCGGTACCGTCTGCTGAGAGACACATCTCTG 3580
Qy 574 CCAGCATGTCGCGCGCTCATCTGCTGACCTGGAACCGCGGAGGAGCGATCGGG 633
Db 3581 CCGGCATCGCGTTCCTCCGTGCTGGCGCGCTCGATCTCTCCGAGGAGCGGAGCA 3640
Qy 634 AGGTGCGCGCG 646
Db 3641 CGCCTCGCGCGC 3653

RESULT 5
LOCUS SC0939127/c 290850 bp DNA linear BCT 11-FEB-2003
DEFINITION Streptomyces coelicolor A3(2) complete genome; segment 24/29.
ACCESSION AL939127 AL023496 AL023861 AL031155 AL031182 AL031232 AL035161
AL035205 AL035206 AL132824 AL512902 AL590982 AL591083 AL591084
AL645882
VERSION AL939127.1 GI:24429552
SOURCE Streptomyces coelicolor A3(2)
ORGANISM Streptomyces coelicolor A3(2)
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1
AUTHORS Bentley,S.D., Chater,K.F., Cerdeno-Tarraga,A.M., Challis,G.L.,
Thomson,N.R., James,K.D., Harris,D.E., Quail,M.A., Kieser,H.,
Harper,D., Bateman,A., Brown,S., Chandra,G., Chen,C.W., Collins,M.,
Cronin,A., Fraser,A., Goble,A., Hidalgo,J., Hornsby,T., Howarth,S.,
Huang,C.H., Kiser,T., Lark,L., Murphy,L., Oliver,K., O'Neill,S.,
Rabinowitsch,E., Rajandream,M.A., Rutherford,K., Rutter,S.,
Seeger,K., Saunders,D., Sharp,S., Squares,R., Squares,S.,
Taylor,K., Warren,T., Wietzorrek,A., Woodward,J., Barrell,B.G.,
Parkhill,J. and Hopwood,D.A.
Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)
JOURNAL Nature 417 (6885), 141-147 (2002)
MEDLINE 21996410
PUBMED 1200953
REFERENCE 2 (bases 1 to 290850)
AUTHORS Bentley,S.D.
JOURNAL Direct Submission
TITLE Submitted (09-MAY-2002) Submitted on behalf of the Streptomyces
sequencing team, Sanger Institute, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: sdb@sanger.ac.uk
COMMENT On or before Oct 30, 2002 this sequence version replaced
gi:20520661, gi:20520749, gi:20520754, gi:20520755, gi:20520688,
gi:20520689, gi:20520815, gi:20520683, gi:20520917, gi:20520866,
gi:20520864.
FEATURES
source Location/Qualifiers
1. .290850
/organism="Streptomyces coelicolor A3(2)"
/mol_type="genomic DNA"
/strain="A3(2)"
/db_xref="taxon:100226"
97. .1101
/gene="SCO6257"
97. .1101
/notes="synonym: SCAH10.22"
97. .1101
/gene="SCO6257"

/note="SCAH10.22, probable ABC transport system sugar
binding lipoprotein, len: 334 aa; similar to TR:CAB41563
(EMBL:AL049727) Streptomyces coelicolor putative secreted
solute binding protein, 337 aa; fasta scores: opt: 1003
z-score: 1130.9 E(0); 49.8% identity in 325 aa overlap,
to SW:R89B_BACSU (EMBL:Z92953) Bacillus subtilis
D-ribose-binding protein precursor RbsB, 305 aa; fasta
scores: opt: 206 z-score: 238.7 E(0); 6.3e-06; 23.0%
identity in 235 aa overlap and to TR:CAB41563 (EMBL:SC981)
Streptomyces coelicolor SC981.17c, 337 aa; fasta scores:
opt: 1003 z-score: 1045.2 E(0); 49.8% identity in 325 aa
overlap. Contains match to Pfam entry PF00532
Peripla_Bp_like, periplasmic binding proteins and LacI
family and a match situated in the correct position to
Prositate entry PS00013 Prokaryotic membrane lipoprotein
lipid attachment site. Contains also possible N-terminal
region signal peptide sequence"
/codon_start=1
/transl_table=11
/product="putative ABC transport system sugar binding
lipoprotein"
/protein_id="CAB60176.1"
/db_xref="GI:6273662"
/translation="MARPTWVGIALGALNSVSLAGCSSTGCKRAEDAKAASAGRA
ADTPRTWPMITHSGDGTDFWDIVQSGAEQAQVNDNINFLSHDRAQQOQLVNAA
IDKKVDGIIVTLAKDPAMKALARAHAQIPVITVNSGSESKFGLTHVGQDETIA
GEAVGEELNREGQAVCVLHQGNVHEQSCDVEKTFDGKVRQLYVNGTSPDVQS
ATEALQTKSDVAVVTLGAPYADTAVAKQAGKAEIDTDLNAKVAAGLADGTLG
FAVDQDPYLGVEAVDLEWLKYNADVLLGGRPVLTGPQIITKDDAAALADYTKRTR
"
133. .165
/gene="SCO6257"
/note="PS00013 Prokaryotic membrane lipoprotein lipid
attachment site"
241. .966
/gene="SCO6257"
/note="Pfam match to entry PF00532 Peripla_Bp_like,
Periplasmic binding proteins and LacI family., score
13.80, E-value 2.1e-06"
1098. .2138
/gene="SCO6258"
/note="synonym: SCAH10.23"
1098. .2138
/gene="SCO6258"
/note="SCAH10.23, probable ABC transport system sugar
permease, len: 346 aa; similar to TR:O68120
(EMBL:AF010496) Rhodobacter capsulatus ribose transport
system permease protein RbsC, fasta scores: opt: 445
z-score: 508.0 E(0); 6.4e-21; 30.5% identity in 351 aa
overlap and to SW:R89B_ECOLI (EMBL:L10328) Escherichia
coli ribose transport system permease protein RbsC, len:
321 aa; fasta scores: opt: 286 z-score: 329.9 E(0);
5.3e-11; 30.1% identity in 329 aa overlap. Contains
possible hydrophobic membrane spanning regions"
/codon_start=1
/transl_table=11
/product="putative ABC transport system sugar permease"
/protein_id="CAB60177.1"
/db_xref="GI:6273663"
/db_xref="GOA:Q9RKT4"
/db_xref="UniProt/TREMBL:Q9RKT4"
/translation="MNAIQVKAADRIILQTSRLKRLGLRPELGSVVGAIAVFFVFAFF
ADSFLLHAASLTSLVYAASTGIMAVFALLMGFFDLSAGVMTSSALVSMTFSYQM
TANVWVGVVSLVTLAIGAFNGFMLTRTKLPSFTITGLFMLGLNGLKLVNDGT
VSTKSIADMEGPFSAQDVFASTITIGGVGFKYTLWLALVAVASWILLRTRAGNIWIF
AVGKNDAARAVGVPAVKTKGLYMGVGVGAWISQHLLEFSDYDVVQSGVSGNELIYI
TAAVIGGCLITGTGYSAGVSAVGAFIFGTSKGIIVFAEWNPDWFRFFLGLMLLATLL
NAWVRKRAEATK"
2124. .2127
/gene="SCO6258"
2135. .2926
/gene="SCO6259"

misc_feature

misc_feature

gene

CDS

RBS

gene

CDS	/note="synonym: SCAH10.24" 2135..2926 /gene="SCO6259" /note="SCAH10.24, probable ABC sugar transport ATP binding protein, len: 263 aa; similar to TR:O68121 (EMBL:AF010496) Rhodobacter capsulatus ribose transport ATP-binding protein, 305 aa; fasta scores: opt: 612 z-score: 706.1 E(): 5.9e-32; 45.6% identity in 217 aa overlap, to SW:RBSA_ECOLI (EMBL:M13169) Escherichia coli, ribose transport ATP-binding protein RbsA, 501 aa; fasta scores: opt: 509 z-score: 585.6 E(): 3e-25; 31.5% identity in 254 aa overlap and to TR:CAAL5787 (EMBL:SC7B7) Streptomyces coelicolor St7B7.07, 260 aa; fasta scores: opt: 592 z-score: 521.2 E(): 1e-23; 43.1% identity in 246 aa overlap. Contains Pfam match to entry PF00005 ABC tran, ABC transporter and Prosite PS00017 ATP/GTP-binding site motif A (P-loop) and PS00211 ABC transporters family signature" /codon_start=1 /transl_table=11 /product="probable ABC sugar transport ATP binding protein" /protein_id="CAB60178.1" /db_xref="GI:6273664" /db_xref="GOA:Q9RK13" /db_xref="UniProt/TREMBL:Q9RK13" /translation="MTRNEDRTALVELSGVSKNYGNVRALEGVSLVHAGETTCVLGD NGAGSTLKIITAGLHQHDGTLSDIDETRLSSPREALDRGIATVYQDLAVVPLMPV WRNPLGSEPRKGVAPFRKMDVDMRRRTTHAELLRMGIDRLVDOPIGTLSGGERQCV AIARAVYFGAKVLIDDEPTAALGVKQGVLYKVAARDQGLGVLIITHNPHAYLVG DRFVLKRGTVANQTRDVEITDELITQQMAGTDELDDLREHERG" 2240..2815 /gene="SCO6259" /note="Pfam match to entry PF00005 ABC tran, ABC transporter, score 168.70, E-value 9.6e-47" 2261..2284 /gene="SCO6259" /note="PS00017 ATP/grp-binding site motif A (P-loop)" 2588..2632 /gene="SCO6259" /note="SCO00211 ABC transporters family signature" 3042..4190 /gene="SCO6260" /note="synonym: SCAH10.25" 3042..4190 /gene="SCO6260" /note="SCAH10.25, possible sugar kinase, len: 382 aa; similar to SW:GLK_STRCO (EMBL:X65932) Streptomyces coelicolor glucokinase (EC 2.7.1.2) (glucose kinase) Glk, 317 aa; fasta scores: opt: 1022 z-score: 1142.9 E(): 0; 46.6% identity in 311 aa overlap and to TR:CAB51974 EMBL:SC6E10 Streptomyces coelicolor SC6E10.20c, 317 aa; fasta scores: opt: 1022 z-score: 1035.4 E(): 0; 46.6% identity in 311 aa overlap. Contains a match to Pfam entry PF00480 ROK, ROK family and Prosite entry PS01125 ROK family signature" /codon_start=1 /transl_table=11 /product="putative sugar kinase" /protein_id="CAB60179.1" /db_xref="GI:6273665" /db_xref="GOA:Q9RK13" /db_xref="UniProt/TREMBL:Q9RK13" /translation="MSTVRDEFTAPIGSRAPVLRTVGTRERRSHLTAPRPVTVGIDIG GTKVADNRVVDAGNLEIKLRTETPDKSPKVEDTIVELVLDLSDRDHVAVGIAA GVVDADRNLVAPILSWRNEPLRIAGRLAVPVLVDNDANTAAWAWRFAGRGG HLVTITLGTGIGGAILEDQVKRGVAGRGHVPVPGHRCPCNGRNGWEYSSG NALVREARLAADSPVAYGIIIEHVKSGIDITGPMIELAREGAMCVELLQIDTQW LGVGNLAALADPSCFVIGGVSAADDLLIGFAPDAFKRLTGTGRGYPEARIVRAQL GPAGVNGAADLSRLVARRFRKRRVRRYERYAEARESESL" 3159..3704 /gene="SCO6260" /note="Pfam match to entry PF00480 ROK, ROK family, score	misc_feature	215.50, E-value 5.4e-63" 3540..3623 /genes="SCO6260" /notes="PS01125 ROK family signature" 4244..4247 4259..4768 /gene="SCO6261"	misc_feature	Query Match Best Local Similarity 32.8%; Score 229; DB 1; Length 290850; Matches 373; Conservative 0; Mismatches 240; Indels 0; Gaps 0;	QY	34	CAGGAACGGCGCGTCCGACCGGCGGCGATCGTGGCGGACGCGCTCTGGTCTTCGAC	93
					Db	11966	CAGGACCGGCGATCCGACGCGGCGAGATCTCTGGAGCGCGCGCAGGTCTTCGAG	11907	
					QY	94	GAGTACGGGTTTCAGGCGCGCCACAGTGGCAGAGATCTCTCGGGGGGCTCGGTCAACAAG	153	
					Db	11906	AAGCAGGGGTACCAAGCTGCCACGATCAGGAGATCTCTCAAGTGGCGGGTGACCAAG	11847	
					QY	154	GGCGCGATGTACTTCCACTTCGCTTCAAGGAAGAGCTGGCCCGCGGTGCTGGCCGAG	213	
					Db	11846	GGAGCGCTCTACTTCCACTTCCAGTCCAAAGGAAGAACTGGCGCTGGCGCGTCTTCGACGCC	11787	
					QY	214	CAGACCTGCGACGTGGCGGTGCCGAATCCGGCTCCAAAGCGCGAGAACTGGTAGACCTC	273	
					Db	11786	CAGGAACACACACAGCGCGTTCGGGAGCAACCCCTCGGCTGCAAGAACTCATCGACATG	11727	
					QY	274	ACATGCTGGTCCGCCACCGGATGCTGCAGATCCGATCCGATCCGCGCGGCGGCAACGCGCTC	333	
					Db	11726	GGCATGTTGTTCTGTACCGCTTGGCGCAAGAACTGCTGGCGCGCGCGCGTGGCGCTC	11667	
					QY	334	GCACTCGACACGCGGGCGGTGGACTTCTCCGACGCCAACCCGTTCCGCGAGTGGGGCGAC	393	
					Db	11666	TCCATGGACACGAGCGCGACGCTCGATCCGCGAGGACCTTCGCTCGCTCGCGCACGAG	11607	
					QY	394	ATCTGCGCCAGCTCTCTGGCGGAGGACAGGAAACGGGGGAGTGTCTTCGCGACGTGAAC	453	
					Db	11606	ACACTCTCTGAAGCTGTGTGAACAGGCGCAAGGAGAACGCTGAGTTGCTGCCCATGTGGTC	11547	
					QY	454	CGMAAAAGACCGGCGACTTCTATCGTGGCTGCTTCCACCGGGCTCCAGCGGTCTCCCGG	513	
					Db	11546	ACCACGACTCGGCGCATCTCTACGTGGGACGCTTCGCGGGGATACAGTGTGTGCCAG	11487	
					QY	514	GTCACCTCCGACCGCGCAGGACCTTCGGCCACCGGATCTCGGTGATGTGGAACACCATGCTG	573	
					Db	11486	ACGGTCAGCGACTTACCAGGACCTTCGAAACACCGCTAGCGGCTGCTGCAGAAACATCTG	11427	
					QY	574	CCGACGATCGTCCGCGCTCCATGCTGACTGATCGAAACCGGCGGAGGCGGATCGGG	633	
					Db	11426	CCCGCATCGCGGTTCCTCCGCTGCTGCGCGCGCTCGATCTCTCCGAGGAGCGCGGAGCA	11367	
					QY	634	AAGGTTCGCGCGCGG 646		
					Db	11366	CGCTTCGCGCGCGG 11354		
RESULT 6									
STWAFR									
LOCUS	STWAFR	1328 bp	DNA	linear	BCT 10-FEB-1999				
DEFINITION	Streptomyces griseus DNA for A-factor receptor, complete cds.								
ACCESSION	D49782								
VERSION	D49782.1	GI:11313911							
KEYWORDS	A-factor receptor.								
SOURCE	Streptomyces griseus								
ORGANISM	Streptomyces griseus								
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;								
AUTHORS	1 (bases 1 to 1328)								
TITLE	Streptomycetaceae; Streptomyces.								
JOURNAL	Cloning and characterization of the A-factor receptor gene from Streptomyces griseus								
	J. Bacteriol. 177 (21), 6083-6092 (1995)								

```

MEDLINE      96042082
PUBMED      7592371
REFERENCE   2 (bases 1 to 1328)
AUTHORS     Onaka,H.
TITLE       Direct Submission
JOURNAL     Submitted (18-MAR-1995) Hiroyasu Onaka. The University of Tokyo,
Department of Biotechnology; Yayoi, Bunkyo-ku, Tokyo 113, Japan
(E-mail: aas7093@hongo.ecc.u-tokyo.ac.jp. Tel:03-3812-2111(ex.5147),
Fax:03-5802-2931)
COMMENT     On May 11, 1996 this sequence version replaced gi:1089808.
FEATURES    Location/Qualifiers
             source          1..1328
                        /organism="Streptomyces griseus"
                        /mol_type="genomic DNA"
                        /strain="IFO13350"
                        /db_xref="taxon:1911"
                        299..1129
                        /codon_start=1
                        /transl_table=11
                        /product="A-factor receptor"
                        /protein_id="BAA08617.1"
                        /db_xref="GI:11742915"
                        /translation="MAKQARAVOTWRSIVDAASVFDYGVRAAISEILRRRAKVTGR
ALVFPASKAIAQA INDEQTSVFEQEGSPQLSLVDGQQAIPALRHNSMARAGTR
LSLEGVLPFGHPGWDITARMLEQGEVFPQIDPMVSAKIVASFITGIQLIVS
EADSGRADLREQVAEMWRHLLPSIAHEGVIAHKPEGRVDLAAQAKAEKAEKREQEAR
AABAAGKAGSDPTSEGGTRSGSLRGSGGRPRAGVTGDEGDEEPAGAAAGGIVA
"
ORIGIN
Query Match      30.6%; Score 213.6; DB 1; Length 1328;
Best Local Similarity 58.7%; Pred. No. 8.1e-17;
Matches 391; Conservative 0; Mismatches 269; Indels 6; Gaps 1;

QY  34  CAGGAAACGGCGCTCCGACGCGGACGGCGATCGTGGCGGAGCCGCGCTCGTCTTCGAC 93
DB  308 CAGCTTCGCGCAGTCCAGACGTGGCGGTGCATCGTGGATGCCGCGGCGAGTGTCTTCGAC 367

QY  94  GAGTACGGGTTCGAGGCGCGCACAGTGGCAGAGATCCCTCTCGCGGGGCTCGGTCAACCAG 153
DB  368 GACTACGGCTACGAGCGTGC CGCCATCTCGGAGATTTCTGCCCGCGCCAGGTCAACCAG 427

QY  154 GGGCGGATGATCTTCCACTTCGCTTCCAAAGAAAGAGCTGGCCCGCGCGCTGCTGGCGCGAG 213
DB  428 GGGGCGCTTGACTTCCACTTCGCTTCCAAAGAGGCCATCGCCCAAGCGGATCATCGACGAG 487

QY  214 CAGACCTTGACGTGGCGGTGCCGGATCCGGCTCAAGGGGCGAGAACTGGTAGACCTC 273
DB  488 CAGACGCTCCACGCTGGAGTTCGAGCAGGAGGGCTCGCCGCTTCAGTCCCTCGTGGACGGG 547

QY  274 ACCATGTGTGTGGCCACGGCACTCTGCACGATCCGATCTCTGCGGGGGCGGACGCGGCTC 333
DB  548 GGCACAGAGTTCGTTTTCGCTTCGGCCACAACTCGATGGCCCGGGCCGGTACCAAGGCTC 607

QY  334 GCACCTGACAGGGGGCGGTGGACTTCTCCGACCGCAACCCCGTTTCGGCGAGTGGGGCGAC 393
DB  608 TCCATCGA-----GGGCGCTTCTCTCGGGGGCGCACCCCTGGGCGCACTGGATCGAC 661

QY  394 ATCTGCCGCCAGCTCTTCGGCGAGGACACAGAAACGGGGGAGGTGCTTCGCGCAGTGAAC 453
DB  662 GCGACGCGCCGGATGCTGGAGCTGGGCGCAGAGCGCGCGAGGTGTTCCTCCGACAGATCGAC 721

QY  454 CCGAAAAGACCGCGCACTTCATCGTCGGCTTCATCCCGGGCTCCAGGGCGGTCTCCCGG 513
DB  722 CCGATGTGTGACGCAAAATCATCGTCGTTCTGTTCAACCGGTATCCAGCTGGTCTCGGAG 781

QY  514 GTCACTCTCCAGCCGACGAGCCTCGGCCACCGGATCTCGGTGATGTGGAAACCAAGTCTG 573
DB  782 GCCGATTCGGCCCGGGCCGATCTCCGCGACGAGTGGCGGAGATGTGCGGCGCATCTCTG 841

QY  574 CCCAGCATCTGTCGGCGGTCCATGCTGACTTGGATCGAAACCGGCGAGGACGCGATCGGG 633
DB  842 CCGTTCGATCGCCACCCCGGGGTTCATCGCCACATCAAAACCGAGGGCCCGGTCGATCTG 901

```

Qy	634	AAGTCCGCGGGCGGCCGAGAGCGCTGCAGAGCCTTCGAGGCGCGCTCCGAC	693
Dd	902	GCGGCCAGGCGCGCGAAGGCGGCGGTGAGGAGCAGGATCGCGCGGAG	961
Qy	694	GAGTAG	699
Dd	962	GCCNAG	967

RESULT 7	AV338477	33064 bp	DNA	linear	HTG 07-OCT-2004
LOCUS	AV338477/c				
DEFINITION	Streptomyces ambofaciens clone cosmid F6 strain ATCC 23877, *** SEQUENCING IN PROGRESS ***, 2 ordered pieces.				
ACCESSION	AY338477	AY338478			
VERSION	AY338477.2	GI:53851167			
KEYWORDS	HTG; HTGS PHASE2.				
SOURCE	Streptomyces ambofaciens				
ORGANISM	Streptomyces ambofaciens				
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.				
AUTHORS	1 (bases 1 to 33064) Pang,X., Aigle,B., Gizardet,J.M., Mangenot,S., Pernodet,J.L., Decaris,B. and Leblond,P.				
TITLE	Functional Angucycline-Like Antibiotic Gene Cluster in the Terminal Inverted Repeats of the Streptomyces ambofaciens Linear Chromosome Antimicrob. Agents Chemother. 48 (2), 575-588 (2004)				
JOURNAL	14742212				
PUBMED	2 (bases 1 to 33064)				
REFERENCE	Aigle,B.				
AUTHORS	Direct Submission				
TITLE	Submitted (08-JUL-2003) Genetique et Microbiologie, UMR INRA 1128, IFR 110, Universite Henri Poincare, Faculte des Sciences et Techniques, Boulevard des Aiguillettes, BP 239, Vandoeuvre-lès-Nancy 54 506, France				
JOURNAL	3 (bases 1 to 33064)				
REFERENCE	Aigle,B.				
AUTHORS	Direct Submission				
TITLE	Submitted (06-OCT-2004) Genetique et Microbiologie, UMR INRA 1128, IFR 110, Universite Henri Poincare, Faculte des Sciences et Techniques, Boulevard des Aiguillettes, BP 239, Vandoeuvre-lès-Nancy 54 506, France				
JOURNAL	Sequence update by submitter				
REMARK	On Oct 7, 2004 this sequence version replaced gi:39748104.				
COMMENT	* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes * of the gaps between them are based on estimates that have * provided by the submittor. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. * 1 13715: contig of 13715 bp in length * 13716 13815: gap of unknown length * 13816 33064: contig of 19249 bp in length. Location/Qualifiers 1. .33064 /organism="Streptomyces ambofaciens" /mol_type="genomic DNA" /strain="ATCC 23877" /db_xref="ATCC:23877" /db_xref="taxon:1889" /clone="cosmid F6" 239..1324 /gene="alpH" 239..1324 /gene="alpH" /note="similar to Pfam PF00891 O-methyltransferase" /codon_start=1 /evidence=not experimental /transl_table=11				
FEATURES	source				
gene					
CDS					


```
Db 32058 CGGCAGACGACGATCCGTACCCAGTTGTCATCTCTGAGGCTGGCGCGAGGTGTTTC 32799
Qy 91 GACGAGTACGGTTTCAGAGCGCCGACAGTGGCAGAGATCTCTTCGCGGGCTCGGTCAACC 150
Db 32798 GACTCCCATGGCTACGAGGCGCGACCACTCGGGAGATCTCTAGACGGCGGGTGTCAACC 32739
Qy 151 AAGGGCGGATGTACTTCCACTTCGCTTCCAAAGAGAGAGTGGCCGCGCGGTCTGGGCC 210
Db 32738 AAGGGCGGCTGTATTTCCACTTCCTTCCAAAGAGAGTGGCCGCGGGTGTGGAG 32679
Qy 211 GACGACAGCTGTCACGTCGGCGGTGCGCGAATCCGGCTTCCAAAGCGCAGGAACCTGGTAGAC 270
Db 32678 CACGAGTTCTCCGTGATCCGGGTGCGCGCGGTCCGTGAGGCTCCAGAGTTCGTGGAC 32619
Qy 271 CTCACCATGTCGTGCGCCACGAGATGCTGCACGATCCGATCTTCGGCGCGGCACCGCG 330
Db 32618 ACGGGTCTGATCTGGCTTACCGCATCGCGCGTGTATCCGTCGGTGGAGCGCGTCCGACG 32559
Qy 331 CTCGCACTGACAGGCGGCGGTGGACTTCTCCGACGCCAACCCGTTCCGCGAGTGGGC 390
Db 32558 CTGAGTCTGGAGCAGGAATGCGTGGGAGTACGGCTGTCGGCGATCCGGCAGTGGATC 32499
Qy 391 GACATCTGCCGCCAGCTCTTGGCGGAGGACACAGAACCGGGGAGGTGCTTCCGACGCTG 450
Db 32498 GGGCGCTCCGAGTGTCTGGGTGCGCGAAGGACGAGGTGAGCTTCTTCGCGATGTG 32439
Qy 451 AACCCGAAAGACCGCGCATCTTCATGTCGCTGCTTCAACCGGCTCCAGGCGGCTCC 510
Db 32438 GTCCGCGCGAGAGTGGTGGCTTCTCCGCGCGTGGAGCGCGTACGACGCTCTACTCG 32379
Qy 511 CGGCTCACCTCCGACCGCAGGACCTTCCGCCACCGATCTCGTGTATGTGGACACGCTG 570
Db 32378 CAGATCTGCTGGCGCGGAGGACCTTGGAGGACCGGTGGTGTCTTCTTCGCGACCTG 32319
Qy 571 CTCGCCAGCATCTGCGCGCTCATGCTGACCTGATCGAATCGAGGAGCGGATC 630
Db 32318 CTCGCCAGCATCTGCGCTGCGCGCTGCTCAGCAGTCTGGAGATCACCCAGCGGGCG 32259
Qy 631 GGAAGGTCCGCGCGCGCGGAGGCGCGGAGCTCGGAGGCTCCGAGGCGCGCTCC 690
Db 32258 GCCCGCTCGTTCGGCGAGCGCGGCTTCTCGAGGCGCGCGCGGAGCGGAGCTCC 32199
Qy 691 GACGAGTAG 699
Db 32198 CGCAGGAG 32190

RESULT 8
AB088224/c
LOCUS Streptomyces rochei plasmid pSLA2-L DNA linear BCT 11-JUN-2003
DEFINITION Streptomyces rochei plasmid pSLA2-L DNA, complete sequence.
ACCESSION AB088224
VERSION AB088224.1 GI:30698345
KEYWORDS
SOURCE Streptomyces rochei
ORGANISM Streptomyces rochei
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE
1 Kinashi, H., Sugino, H., Sasaoka, A., Mori, E., Fujii, S., Shinkawa, H.,
Nimi, O. and Kinashi, H.
Physical mapping of the linear plasmid pSLA2-L and localization of
the eryA and actI homologs
Biosci. Biotech. Biochem. 62, 1992-1897 (1998)
2 Suwa, M., Sugino, H., Sasaoka, A., Mori, E., Fujii, S., Shinkawa, H.,
Nimi, O. and Kinashi, H.
Identification of two polyketide synthase gene clusters on the
linear plasmid pSLA2-L in Streptomyces rochei
Gene 246 (1-2), 123-131 (2000)
JOURNAL 20231737
MEDLINE 10767533
PUBMED
AUTHORS
REFERENCE 3
Hiratsugu, K., Mochizuki, S. and Kinashi, H.

Cloning and analysis of the replication origin and the telomeres of
the large linear plasmid pSLA2-L in Streptomyces rochei
Mol. Gen. Genet. 263 (6), 1015-1021 (2000)
20408175
PUBMED 10954087
REFERENCE 4
Mochizuki, S., Hiratsugu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K.
and Kinashi, H.
The large linear plasmid pSLA2-L of Streptomyces rochei has an
unusually condensed gene organization for secondary metabolism
Mol. Microbiol. 48 (6), 1501-1510 (2003)
22676866
PUBMED 12791134
5 (bases 1 to 210614)
Mochizuki, S., Hiratsugu, K. and Kinashi, H.
Direct Submission
Submitted (15-JUL-2002) Haruyasu Kinashi, Hiroshima University,
Department of Molecular Biotechnology, Graduate School of Advanced
Sciences of Matter; 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima
739-8530, Japan (E-mail: kinashi@hiroshima-u.ac.jp,
Tel: 81-824-24-7869, Fax: 81-824-24-7869)
The nucleotide sequence has been determined by using restriction
fragments and nested deletion fragments of the ordered cosmid
library of pSLA2-L.
pSLA2-L is 210,614 bp long with G+C content of 72.8 % and carries
143 ORFs.
Gene prediction was based on the unique codon usage in Streptomyces
(Bibb et al., Gene 30:157-66 (1984)) using the FramePlot program of
Ishikawa and Hotta (FEMS Microbiol Lett 174: 251-253 (1999)) as
implemented at
http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl. Where possible we
chose an initiation codon (atg, gtg, ctg or ttg) which is preceded
by an upstream ribosome binding site sequence (optimally 5-13 bp
before the initiation codon). If this could not be identified we
chose the most upstream initiation codon.
Three-fourths of the pSLA2-L DNA are occupied by biosynthetic genes
(three PKS gene clusters for lankacidin, lankamycin and an unknown
type II polyketide, a carotenoid biosynthetic gene cluster, many
regulatory genes and others). The range of each biosynthetic gene
cluster has been deduced by comparison with similar gene clusters
in most cases and may be revised in future.
Location/Qualifiers
1. .210614
/organism="Streptomyces rochei"
/mol_type="genomic DNA"
/strain="7434AN4"
/db_xref="taxon:1928"
/plasmid="pSLA2-L"
/note="linear plasmid"
1. .1992
/note="left terminal inverted repeat, TIR-L; shows 99.4 %
(1981/1992) sequence identity to TIR-R (complement
(208623..210614))"
683..2188
/note="N-terminal sequence is almost identical (435/437)
with that of ORF143 at the right end of pSLA2-L until the
inner end of TIR-L
ORF1 (501 aa)
similar to AL590463 Streptomyces coelicolor putative
helicase, SCP1.136 (879 aa); homology is seen until the
inner end of TIR-L"
/codon_start=1
/transl_table=11
/product="putative helicase"
/protein_id="BAC76459.1"
/db_xref="GI:30698346"
/db_translation="MSTTSRTDREAQQEAVDAVVRALLEPVRLLAPRGLRTQVIM
ATGSGKTRVAARSARKRAGRVLLVPSLDLTLQTEAAWREAGRTGPMIGVSLRGED
VAFPNITDVEELVDWVRPFDKTVTFATYASGLGLTERAHGGLPGWDLIVVDEAHT
SGRLGKPAVVDHNTRIPSRLRYMTATPRMLQDLQDAECAPGELVASMEDPDGLFG
ARCFTLTSEAIDRGICAPYQVVDITDLOQAQLGVEGRSDVRGRLAALQTA
LLKASSENFRETLVPHNVKEAEAPAGLPDVAKRLHAAGPQLVPRTTWVWLCSEH
KFGHRRRVLGFTSGIATDGTVEKGFVGSLGVEGVDTRCEDSVYADVRGSMFPL
```

CDS

VOAGRALRIPQGGKVASLVVPLDPEETADNMLTGRPYNGLALLFQAPLLTGHG
 ESGFAARPPGVEFVLCVGRGAQAGAPASRSITPSAMSSLSISWRAP"
 /complement (2315..3595)
 /note="ORF2 (426 aa)
 similar to AE004736-10 Pseudomonas aeruginosa hypothetical
 protein (442 aa)"
 /codon_start=1
 /transl_table=11
 /product="putative secreted protein"
 /protein_id="BAC76460.1"
 /db_xref="GI:30698347"
 /translation="MRKRALLGILPLGLLVAVLAPAAHAASADTYVGTWAPPTAA
 PAADSTVEDQTLRQVHLSVAGSLRVRFTNEFGTSLPTIGEVHAARPAACGPATV
 DPGTRVRFGRGPATLPGTQRMWDVALPTAGGLVLSLYLPQTGSGTSHAA
 YGHNVAAGDVGTGADPLTPVSTATSMHFLSGVSDRRAGTADSSVVTGLDSTIDGHT
 TLDAERVDLLAELRLDRDGLAGTGVNAGIGGNLLLRDPDPEGSAAESFAAYEGE
 SALKFRDRLQPGARAVTVLLGVNDLQPGIAPASDEVTAABELIAGIQLIEAH
 EHLKIDRTTFFAGDTTGYFTPREAVRQVNDWIRTSGAFDVTLPDFAVLDRPAR
 PDHLLPAYDGGDLHPNDAGMAARAFDLRLR"
 /complement (3659..4645)
 /note="ORF3 (328 aa)
 similar to Y00459-2 Streptomyces griseus regulatory
 protein, StrR (350 aa)"
 /codon_start=1
 /transl_table=11
 /product="StrR-like regulatory protein"
 /protein_id="BAC76461.1"
 /db_xref="GI:30698348"
 /translation="MTAALAGNLQDQETRIPISLLVPGYSPRLCGEDAEHVARLADTD
 EALPFLVERSLRVIDGHRVLAQAKGHITIEVLFDFGAEEAFLLAVRNMTHGL
 PLRQDRRAAQRILQAQPHLSDRVAGIAGAKTAAALRPLAAGPTNPQARGRD
 GRIRLDGTTGRKQAEALRQPAQSVREVARHAGISPATASDVRRLASGRSPYPER
 NTPGARPGCTGSRATPGGGAAPVPSVVRPPEDSPVRLRLDPSLRHESGRRL
 LLQCGAVERTALLAMACTVPPHCTDLVAELAREYADLWAEFAREVARTDG"
 /complement (4922..5812)
 /note="ORF4 (296 aa), lankacidin biosynthesis protein
 similar to AE004625-4 Pseudomonas aeruginosa
 pyrroloquinoline quinone biosynthesis protein B (304 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein B"
 /protein_id="BAC76462.1"
 /db_xref="GI:30698349"
 /translation="MILLTAAAGGFPFQWNCALCARGRRGELPARSOECVAVSGDG
 RWMNLNAGSDRTQLLAAPALTPGPRDTPVRGVLLTDAEDVHALGLVLRGATGL
 TVVAPVPVAGSALAEPLVRGLDRDPAWDRTATPGGFAVAGGLTVTAHPVGTKAPK
 YAHAPDPAPWCAYRIEDPATGGALVYAPCLATWPDGFDLLASATCALLDGTFFSA
 GELGTATSAAGAQSLMGLHPVAGPGSLLAALRHRLRIYTHLNTNPLDPPSAA
 HAAVRAGVEVLPDGSSELVL"
 /complement (5815..6909)
 /note="ORF5 (364 aa), lankacidin biosynthesis protein
 similar to AJ277117-5 Gluconobacter oxydans putative pqqE
 protein (359 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein E"
 /protein_id="BAC76463.1"
 /db_xref="GI:30698350"
 /translation="MADPAVAGPAGMLIELTHRCPLHCPVCSNPLVRRREALTCQ
 WDTLQARELGVMQVHFGSGEPLAPDPLVGHARRLGAVNVLVTSQVGLTAERAH
 DLARRGVHVQLSLQADPAAGAIAGARVHTAKLEAARAVTAAGLPLTVNLVLRGN
 IDGTGRMDVLADLGDRIELANTQYVGLNRRAALMPTAQAARBAARVARTRY
 AGPELVVYAADYDRIKFCMDQWGTSTQTLVTPAGDVLVPCPAAYAITLPPVENLR
 PLSEIYASRSFNAYRGTMREPCRTCPERHADHGGCRQAFQLTGDAAATDPACGL
 SPHRSVDAALAEVTPGVPVAFVPRGFVPA"
 /complement (6902..7174)
 /note="ORF6 (90 aa), lankacidin biosynthesis protein
 similar to A603642-200 Sinorhizobium meliloti putative
 pyrroloquinoline quinone biosynthesis protein D (98 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein D"
 /protein_id="BAC76464.1"

CDS

/db_xref="GI:30698351"
 /translation="MTGLPEPTVRLRFGVRLTRDPARGELALLPERVVVNDTAAAV
 LAHCDGTTSLAGIVERLAEYEGVSAEDVRELLRLAQRVVVDLHG"
 /complement (7171..7899)
 /note="ORF7 (242 aa), lankacidin biosynthesis protein
 similar to AE004625-5 Pseudomonas aeruginosa
 pyrroloquinoline quinone biosynthesis protein C (250 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein C"
 /protein_id="BAC76465.1"
 /db_xref="GI:30698352"
 /translation="MSMSVTREVAAWPMSEAFRQRLHALESYWRHHPFRHMGELL
 DEGERLWAARWYRCLPQKDAIVANCLPEVRQWLRSIVVHDGADACAGAEK
 WRLAEAVGLARDEVDRLVAGTRFAVDVDFARRPWLAAASGLITELFSPOLL
 AHRQLREHVPWIAEYEGVEFTARIYVVGPEGRSLDLVARHAVRSREQEACVRALA
 PKCYLVNAVLSLDYHTCGATRS"
 /complement (7926..8030)
 /note="ORF8 (34 aa), lankacidin biosynthesis protein
 similar to AL603642-197 Sinorhizobium meliloti putative
 pyrroloquinoline quinone biosynthesis protein A (31 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein A"
 /protein_id="BAC76466.1"
 /db_xref="GI:30698353"
 /translation="MRTSGKELPAKAWHRRPFDVITDGMETVAYFSR"
 8420..10057
 /note="ORF9 (545 aa), possible lankacidin resistance
 protein"

Query Match 30.2%; Score 211.4; DB 1; Length 210614;
 Best Local Similarity 59.9%; Pred. No. 3.5e-17;
 Matches 391; Conservative 0; Mismatches 256; Indels 6; Gaps 2;

QY	34	CAGGACGGCGCTCGGACGGCGGAGCGATCGTGGCGGAGCGCCCTCGTTCGAC	93
DB	143115	CAGGACGGCGGATCGGACCGCGCGCGGTTCGAAAGCGGACCGATTCGCC	143056
QY	94	GAGTACGGGTTCGAGCGGCCACAGTGGCAGAGATCTCTCGCGGCGCTCGTTCACCAAG	153
DB	143055	GAGCAGGGGTACCGCGCGGACCGGTGGCGGACATCTCAAGTGGCGCGCTGACCAAG	142996
QY	154	GGCGGATGATCTTCACTTCCTCCAAAGAAAGAGCTGGCGCGCGCTGTCGGCGGAG	213
DB	142995	GGCGCGCTGATCTTCACTTCCTCCAAAGAAAGCGCTTCGCCGCGGACCTCTGGAGGC	142936
QY	214	CAGACCTCGACGTGGCGGTGCGGAAATCCGGCTCCAAAGCGCGAGAACTGGTAGACCTC	273
DB	142935	CAGGTGCGCGACAGCTGTTTCGCA---GCAGCTCAAGCTCCAGGAATGGGTGGACGG	142879
QY	274	ACCATGCTGTCGCCACCGCATGTCACGATCCGATCCTGGCGGCGGCGACCGCGCTC	333
DB	142878	GGCATGACCTGGCGCACCGATGCGCGGAGCCCGTCTGCGCGCGCGCGCGCGCTC	142819
QY	334	GCATGGACAGGGGGCGGTGACATCTCCGACGCCAACCCGTTCCGGCAGTGGGGCGAC	393
DB	142818	TCGGCGGAGCA---CACCGGAGCGAGCAGCAGCGCGCTTCCCACTCGATCGCC	142762
QY	394	ATCTGCGCCAGCTCTCTCGCGGAGGACAGGAAACGGGGGAGGTGCTTCGCCACGTGAAC	453
DB	142761	TTCTCGGCTCGCTCTTTCAGCAGCGGAGGAGGTCTTGGGCCACATCGAG	142702
QY	454	CCGAAAAAGACCGGCGACTTCATGTCGCGGTCTTCCACCGGGCTCCAGCGGGTCTCCCGG	513
DB	142701	CCGGCGGAGACCGCGAATGCGTCTCGGCTCTCTCCAGGGCATCCAGTCTCTCTCCAG	142642
QY	514	GTCACTCCGACCGCCAGCCTCGGCCACCGCATCTCGGTGATGTGGAAACACACGTGCTG	573
DB	142641	TTGACAGCAATGGCGGACATCGAGCAGCGTCTCGCGCTGTTCGGGACAGCTCTC	142582
QY	574	CCGAGCATGTCGCGCGCTCCATGCTGACCTGGATCGAAACCGGCGGAGCGGATCGGG	633

RESULT 12
AY256849
LOCUS
DEFINITION
Streptomyces ansochromogenes putative gamma-butyrolactone-binding
protein gene, complete cds.
ACCESSION
AY256849
VERSION
AY256849.1
KEYWORDS
GI:32765840
SOURCE
Streptomyces ansochromogenes
ORGANISM
Streptomyces ansochromogenes
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomycetes.
REFERENCE
1 (bases 1 to 901)
AUTHORS
Li, W. and Tan, H.
TITLE
Sabr: an important gene involved in regulating nikkomycin
production in Streptomyces ansochromogenes
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 901)
AUTHORS
Li, W. and Tan, H.
TITLE
Direct Submission
JOURNAL
Submitted (16-MAR-2003) Molecular Microbiology, Zhong Guan Cun Bei
Yi Tiao No.13, Beijing 100080, China
FEATURES
Source
1..901
/organism="Streptomyces ansochromogenes"
/mol_type="genomic DNA"
/db_xref="taxon:115647"
239..901
/codon_start=1
/transl_table=11
/product="putative gamma-butyrolactone-binding protein"
CDS
Query Match 28.0%; Score 196; DB 1; Length 901;
Best Local Similarity 57.9%; Pred. No. 1.2e-14;
Matches 368; Conservative 0; Mismatches 265; Indels 3; Gaps 1;

AY256849 901 bp DNA linear BCT 01-DEC-2003
Streptomyces ansochromogenes putative gamma-butyrolactone-binding
protein gene, complete cds.
AY256849.1 GI:32765840
Streptomyces ansochromogenes
Streptomyces ansochromogenes
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomycetes.
1 (bases 1 to 901)
Li, W. and Tan, H.
Sabr: an important gene involved in regulating nikkomycin
production in Streptomyces ansochromogenes
Unpublished
2 (bases 1 to 901)
Li, W. and Tan, H.
Direct Submission
Submitted (16-MAR-2003) Molecular Microbiology, Zhong Guan Cun Bei
Yi Tiao No.13, Beijing 100080, China
Location/Qualifiers
1..901
/organism="Streptomyces ansochromogenes"
/mol_type="genomic DNA"
/db_xref="taxon:115647"
239..901
/codon_start=1
/transl_table=11
/product="putative gamma-butyrolactone-binding protein"
/protein_id="AAP87374.1"
/db_xref="GI:32765841"
/translation="NAOQDRAVTRRAVLRAAAVFAERGYAAATISEILKRAVTKG
ALYFDFSKAALAOQVLOELTPEYHRELKLOEWVADGALARRLPPEFLAGVR
ISADRPGRDVLGSAWPAWRLTSHALTEAKKEGVLPHVVPETAQVFLGAMVGAQFV
SQTLGAWADLDRTAALVGHLLAAIAAPVLTRLDTAPDRGARVTAERRSDDLGI
AC"
ORIGIN
Query Match 28.0%; Score 196; DB 1; Length 901;
Best Local Similarity 57.9%; Pred. No. 1.2e-14;
Matches 368; Conservative 0; Mismatches 265; Indels 3; Gaps 1;
QY 16 GAACGGGTGGCAGTCCGACAGAAACGGGCCGTCCGACACCGCAGCGGCGGACGATCGTGGCGGA 75
Db 230 GGAGGGTGTGTGGCTCAGCAGACCGAGCGGTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 289
QY 76 GCGGCTCGGTCCGACAGTACGGGTTCGAGGCGCGCCACAGTGGCAGAGATCTCTCG 135
Db 290 GCGGAGCGGTCTTCGCGAGCGCGGATACGCGGCGCGCCACCATATCCGAGATCTCCAAG 349
QY 136 GCGGCTCGGTCCGACAGGCGCGGATGTACTTCCACTTCGCTTCCAAAGGAAGAGCTGGCC 195
Db 350 GCGGCGGTGTCCGACAGGCGCGCGCTGTACTTCCACTTCGACTCCAGGCGCGCGCTCGCC 409
QY 196 GCGGCGGTGTGGCGCAGACCGCTGCACTGGGCGGTGCCGGAATCCGGCTCCAAAGCG 255
Db 410 CAAGGGTGTACAGAAACAGCTGACCGCCGAGTACACCTGCGCGGAGCTGGAAGTTA 469
QY 256 CAGGAACGTGTAGACTCACCATGCTGGTTCGCGCCACCGCATCTGCACCATCGATCCGATCTCG 315
Db 470 CAGGAGTGGGTGGAACGCGGGGATGGCCCTGGCGAGCGGCTGCCCGGGAACCCCTTCCTG 539
QY 316 GCGGCGGGGACCGCGGTCCGACTGGACACGAGGGGGGGTGGAGCTTCTCCGACGCAACCGCG 375
Db 530 CTCGCGGGGTCCGATCTCCCGGACCGCGCGCGCGCGGTGACGTGCTCCGAGCGCGCTGG 589
QY 376 TTCGCGAGTGGGGCGACATCTGCGCCCGCAGCTCTTCGGCGGAGCGCACAGGAAACGCGGGGAG 435
Db 590 CCGGCTGGGCGGAGTGAACCTCGACGCGGCTCA---CGAGGCCAAGAAACGCGGCGAG 646
QY 436 GTGCTTCGACAGTGAACCCGAAAAGACCGGGGACTTCATCGTGGCTGCTTCACCGGG 495

```

Db      647 GTGTCGCGCAGTGTGTCGCCGAGAGACGGCCAGGTGTTCTCGGCGCTTGGTG3GC 706
QY      496 CTCGAGCGGTCTCCGGGTACCTCCGACCCGAGGACCTCGGCCACCGGATCTCGGTG 555
Db      707 GCGCAGTTCGTCTCGCAGACCTCGCCGGCTGGCGGACCTGGAACGACGCGCGCGG 766
QY      556 ATGTGGAACCACTGTCTGCCAGCATCGTCCGGGTCCATCTGACCTGGATCGAAACC 615
Db      767 CTGTAGGCACTCTCTCGCGCGATCGCGGCCCGCCGCTCTCACCGGCTGACACC 826
QY      616 GCGCAGGACCGATCGGGAAGTTCGCGCGCGCGCC 651
Db      827 GCGCGGACCGCGCGCGGTGATCGCGGAGGCC 862

RESULT 13
AF156161
LOCUS      Streptomyces tendae TarA gene, complete cds.          BCT 10-NOV-1999
DEFINITION
ACCESSION AF156161
VERSION   AF156161.1  GI:6318602
KEYWORDS
SOURCE    Streptomyces tendae
ORGANISM  Streptomyces tendae
REFERENCE 1 (bases 1 to 2637)
AUTHORS  Engel, P. and Scharfenstein, L.L.
TITLE    Isolation of a gene specifying a gamma-butyrolactone-binding
JOURNAL  protein from Streptomyces tendae
REFERENCE 2 (bases 1 to 2637)
AUTHORS  Engel, P. and Scharfenstein, L.L.
TITLE    Direct Submission
JOURNAL  Submitted (02-JUN-1999) Food and Feed Safety, SRRC, USDA, 1100
FEATURES  Robert B. Lee Blvd., New Orleans, LA 70124, USA
          Location/Qualifiers
          1..2637
            /organism="Streptomyces tendae"
            /mol_type="genomic DNA"
            /strain="ATCC31160"
            /db_xref="ATCC:31160"
            /db_xref="taxon:1932"
          626
            /note="transposon Tn4560 insertion site"
          1922..2584
            /function="nikkomycin biosynthesis"
            /note="gamma-butyrolactone-binding protein"
            /codon_start=1
            /transl_table=11
            /product="TarA"
            /protein_id="AAF06961.1"
            /db_xref="GI:6318603"
            /translation="MAQQQAVRTRAVLRAAAQVPAERGAAYAAIISILKRAQVTKG
            ALYHFDLSKAAALQVQLPEYHLPRELQEWVDAGMTLARRLPREPFLLAGVR
            ISADREGLVGSANPAWSRLTSHVLEAKRGVLPVHPVEETAQVFLGAVGQFV
            SOTLAGWEDLDRTAALYSHLLGATAAPPVLTPLDTPDRGARVIAEARRRSGDLSGI
            AC"

misc_feature
626
CDS
27.8%; Score 194.4; DB 1; Length 2637;
Best Local Similarity 57.7%; Pred. No. 1.3e-14;
Matches 367; Conservative 0; Mismatches 266; Indels 3; Gaps 1;

QY      16 GAACGGGTGCATGCGACAGGAACGGCGCTCCGACCGCGGCGGATCGTGGCGGCA 75
Db      1913 GGAGGTGTGTGCTGACGAGGACCGGACCGCTCCGACCGCGCGCGCTGTCGCGGCC 1972
QY      76 GCGCGCTCGGTCTTCACACAGTAGTCGGGTTCGAGCGCGCCACACAGTGGCAGAGATCTCTCG 135
Db      1973 GCGGACGGCTCTTCGCGGAGCGCGGATACGCGCGCGCCACCATCTCCGAGATCTCTCAAG 2032

```

```

QY      136 CGGCGCTCGGTACACAAAGGGCGGATGTACTTCCACTTCGTTCACAGGAAGAGTGGCC 195
Db      2033 CGGCGCGGTGTACCAAGGGCGGCTGTACTTCCACTTCGTTCACAGGAAGAGTGGCC 2092
QY      196 CGGCGGTGTGCGGCGGACAGACCTGACCTGCGGTGGGTGCGGAATCGGCTCCAAAGCG 255
Db      2093 CAAGGGGTGTTGCGAGGAACAGCTGACACCGGAGTACACCTGCGCGCGGAGTGAAGTTA 2152
QY      256 CAGGAACCTGGTAGACCTCACCATCTGCTGCTGCCCAACGGCATGTGTCACGATCCGATCCTG 315
Db      2153 CAGGAGTGGGTGAGCGCGGGATGACCTTGCCGACAGCGGTGCCCGGGAACCTTCTCTG 2212
QY      316 CGGCGGCGACCGCGCTCGACCTGACACGAGGGGGGTGGAATCTTCGACGCGCAACCG 375
Db      2213 CTCGCGCGGTCCGGATCTCGCGCGACCGCGCGCGGAGGTGCTCGGACGCGCTGG 2272
QY      376 TTCGCGAGTGGGCGACATCTGCCCGCCAGCTCTTGGCGGAGGACACGAGCGGGGAG 435
Db      2273 CCGGCTGGTCAAGGCTGACCT---CGCACGTGCTACCGAGGCCAAGAGCGGGGCGAG 2329
QY      436 GTGCTTCGCGACGTGAACCCGAAAGAACCGCGCGCATCTTCATCGTCGGTGTTCACCGGG 495
Db      2330 GTGCTGCCGCGACGTGGTCCCGGAGGAGACGCCAGGTGTTCTCGCGCTGGGTGGC 2389
QY      496 CTCGAGCGGTCTCCGGGTCACTCCGACCGCGGAGGACCTCGGCCACCGGATCTCGGTG 555
Db      2390 GCGCAGTTCGTCTCGCAGACCTCTCGCGGCTGGGAAGACCTGGACGACCGCGCGCGG 2449
QY      556 ATGTGGAACCACTGTCTGCCAGCATCGTCCGGGTCCATGCTGACCTGGATCGAAACC 615
Db      2450 CTGTACACGCACTCTCTCGCGGATCGCGCGCGCGCGCTCTCACCGGTGACACC 2509
QY      616 GCGCAGGAGCGGATCGGGAAGTTCGCGCGCGCGCC 651
Db      2510 GCGCGGACCGCGGTGCCCGGTGATCGCGAGGCC 2545

RESULT 14
AF117439/c
LOCUS      Streptomyces carzinostaticus subsp. neocarzinostaticus
DEFINITION
ACCESSION AF117439
VERSION   AF117439.1  GI:28192466
KEYWORDS
SOURCE    Streptomyces carzinostaticus subsp. neocarzinostaticus
ORGANISM  Streptomyces carzinostaticus subsp. neocarzinostaticus
REFERENCE 1 (bases 1 to 92294)
AUTHORS  Liu, W., Nonaka, K., Nie, L., Bae, J., Zazopoulos, E., Farnet, C.M. and
          Shen, B.
TITLE     Biosynthesis of the chromoprotein enediyne neocarzinostatin in
          Streptomyces carzinostaticus: analysis of the gene cluster and
          deduction of the biosynthetic pathway
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 92294)
AUTHORS  Liu, W., Nonaka, K., Nie, L., Bae, J., Zazopoulos, E., Farnet, C.M. and
          Shen, B.
TITLE     Direct Submission
JOURNAL  Submitted (03-JUN-2002) Ecopia Biosciences Inc, 7290 Frederick
          Banting, Saint-Laurent, QC H4S 2A1, Canada
FEATURES  Location/Qualifiers
          1..92294
            /organism="Streptomyces carzinostaticus subsp.
            neocarzinostaticus"
            /mol_type="genomic DNA"
            /strain="ATCC 15944"
            /sub_species="neocarzinostaticus"
            /db_xref="ATCC:15944"
            /db_xref="taxon:167636"
            10097..10894
              /codon_start=1
          CDS

```


/protein_id="AAM77989.1"
/db_xref="GI:28192476"
/translation="MILAGKGTSLRPYTRSLPKPLVPIGDSEFSLILEIVLSQSRQF
DVTYLAIGLHLIRFVGGKRWNNHITVASEDTPRGITGVPLGIDLDQPHFLVMN
GDLLDLDVADLAHTRGEAPLTVATCTREVRIDFGLVLTTSNRVYHVEKPTVDYR
VSGVYGLSRTTLEHTYPSHMFPGFDDLLDLAGVQPAEYFPDGYWLDIGRPDDYDR
ANAEFDSIKGTLLKGA"
complement (27203..28198)
/codon_start=1
/transl_table=11
/product="NDP-hexose-4,6-dehydratase"
/protein_id="AAM77990.1"
/db_xref="GI:28192477"
/translation="MAGLVATGAEGFIGSHLVETLVRRHVRVAVOVNSFASWGL
DLDAEVMHVEFPGDVRDPAVLEVPASVYVHLLAIPYSPYAPRPIYLTNV
MGTNVLEAVRATGTPRLVHTISSETISGARTVPIGEDHPLOQAQPYAAKVAADLV
ESTHLSYGVPATLRFNFTGPRQSAVAIPTIIISOLVAGARQITGLSAPRDFTYV
TDAEAFVGTAPASAVIGFNAGTGTVEIGRLAADIMRLMDADASIRDPQRLR
PKDSEVRLVCDAGSLRARTGWRPRTDREDGLRRTTEWPNPVGAYRQTHYNR"
complement (28341..29060)

Query Match 25.9%; Score 180.8; DB 1; Length 92294;

Best Local Similarity 56.8%; Pred. No. 2.1e-13;

Matches 332; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

QY 21 GGTGGCAGTGGCAGCAGGACGGGCGTCCGACGGCGGAGGCGATCGTGGGCGAGCGCG 80
DB 69432 GCCGGAACCGAGCAGGACGGGCCACGAGGACACGAGGTGTGATCTCGTGGCGCTGC 69373

QY 81 CTCGGTCTTCGACGAGTACGGGTTGAGGCGCGCCACAGTGGCAGAGATCCTTCGCGGCG 140
DB 69372 TGAGGTGTTCGACGAGCGGATACAGCGGGCGAGCATCAGTAAATCATGAACCGCGC 69313

QY 141 CTCGGTCCACAGGGCGGATGTACTTCCACTTCGCTCCAGGAGAGCTGGCGCGCG 200
DB 69312 CGCGGTGACCCAGGGGGGATGTACTTCCACTTCAAGTCGAAGAGAGCTGGCGCTGCGC 69253

QY 201 CBTGCTGGCGGACGACACCTGACGCTGGCGGTGCGGAATCCGGTCCAGCGCGCAGGA 260
DB 69252 AGTGATGACAGTCACGACACCTTCATCGACTTCCCTCAGAGGTGCGGCTGCAGCG 69193

QY 261 ACTGTTAGACTCACCATGCTGTGTCGCCACGCGATGTGCGAGTCGATCGATCGCGGCG 320
DB 69192 AGTGATCGACCTCACCTTCGCTCAGGAGTCCAGAGTCCAGACCAACCCCTGATCGGCG 69133

QY 321 GGCACGCGGCTCGCACTGACAGGGGGGGTGGACTTCTCGAGCGCAACCGTTCGG 380
DB 69132 CAGCATCCGGCTCGCGTCGAGCAGGAGAGTTCGGCGTCCGCGACGACACCGCTTACCA 69073

QY 381 CGAGTGGGCGACATCTGCGGCCAGCTCCTGCGGAGGACAGAACGGGGGAGGTGCT 440
DB 69072 GGAATGGGTAGCCCTGGTGAGACCTACTGCTGGAGGCGGCCAGCGGGCGACTGCT 69013

QY 441 TCGCAGCTGAACCCGAAAGACCGGAGCTTCATGCTGGGTGCTGCTTACCGCGGCTCCA 500
DB 69012 CCGGAGGTGGACGAACAGCACTTCGCCGCTGCTTCGTCAGCGCGCTTCCCGGTAGCCA 68953

QY 501 GCGGCTCTCCGGTTCACCTCCGACCGCCAGACCTTCGGCCACCGGATCTCGGTATG 560
DB 68952 GCTGTTCTGAGGCTCGCCAGGCGCCGCTGCGATCTGCCAGCAGGATCTGCTGCTG 68893

QY 561 GAACCACTGCTGCCAGCATCGTGGCGCGCTGCATGCTGACT 604
DB 68892 GCGTTATCTGCTGCTGCGCTCACGCCCTCCGACGAGGGCT 68849

RESULT 15

ABI06894

LOCUS

DEFINITION Streptomyces clavuligerus scar gene for butyrolactone autoregulator

receptor protein, complete cds.

AB106894

AB106894.1

VERSION GI:293335760

KEYWORDS

SOURCE

ORGANISM

Streptomyces clavuligerus
Streptomyces clavuligerus
Bacteria; Actinobacteridae; Actinobacteriales;
Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE

AUTHORS

Kim, H.S., Lee, Y.J., Lee, C.K., Choi, S.U., Yeo, S.H., Hwang, Y.I.,
Yu, T.S., Kinoshita, H. and Nihira, T.

TITLE

Cloning and characterization of a gene encoding the
gamma-butyrolactone autoregulator receptor from Streptomyces
clavuligerus

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

Streptomyces clavuligerus

Streptomyces clavuligerus

Bacteria; Actinobacteridae; Actinobacteriales;

Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

Streptomyces clavuligerus

Streptomyces clavuligerus

Bacteria; Actinobacteridae; Actinobacteriales;

Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

Streptomyces clavuligerus

Streptomyces clavuligerus

Bacteria; Actinobacteridae; Actinobacteriales;

Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

```
QY 430 GGGAGGTGCTTCGGCACGTGACCCGAAAGACCGGCGACTTCATCGTCGGCTGCTTC 489
Db ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 625 GGGAGGTGCTGCCCCACCGCGATCTGGACAGCGTCTCCGGCTCTTCGCCGCTGTAC 684
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 490 ACCGGCTCCAGCGGTCTCCCGGTCACTCCGACCGCAGGACCTCGGCCACCGGATC 549
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 685 ACCGGATCCAGGTGCTGTCGGATCATGACGACCGCCAGGACCTGGCCGAGCGATG 744
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 550 TCGGTGATGTGGAACCAACGTGTCGCCAGCATCGTCGCCGCGTCCATGCTGACCTGATC 609
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 745 TCGACATGTATCGGAATCTGATGCCCGGCTGGCGTCCCTCCGTGCTGGCGCGCTG 804
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 610 GAAACGGCGAGGAGCGGATCGGGAAGTCCGGCGCGCGCGCGCGCGCGCGCGCG 669
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 805 GACTTCTCGACGAGACCGGGGGCGCTCGTGTACGAGCGGTCTATCGAGCGAGCCAGGAG 864
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 670 GAGGCTCCGAGGCGCGCTCC 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 865 CAGGACGAGGCGCGCGGTAC 885
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: March 6, 2005, 21:14:08
Job time : 4026 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:44:49 ; Search time 588 Seconds
(without alignments)
7037.244 Million cell updates/sec

Title: US-10-049-710A-1

Perfect score: 699

Sequence: 1 atggcagtcgacacgaacg.....agggccgctccgacgagtag 699

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	699	100.0	699	6	AAL44719 S virgini
2	229	32.8	4346	6	Abn86379 Scba, Scb
3	229	32.8	4346	10	Adel4794 Streptomy
c 4	126	18.0	12905	3	Aaz99101 S. fradia
5	126	18.0	12905	3	Aaz99100 S. fradia
c 6	88.6	12.7	1819	8	Abv75877 Streptomy
7	85.8	12.3	9521	4	Aas07627 Streptoco
8	78.2	11.2	4257	2	Aav68520 The nucle
9	78.2	11.2	4257	2	Aav10362 Infected
10	68.2	9.8	1224	8	Ac37971 Prokaryot
c 11	68.2	9.8	32329	12	Ado51695 Streptomy
c 12	67.8	9.7	12001	2	Aa076213 HSV L/ST
13	67.2	9.6	1041	8	Aa54248 Pseudomon
14	67.2	9.6	1041	8	Ac42557 Prokaryot
15	66	9.4	1107	11	Abd05332 Pseudomon
16	66	9.4	1134	11	Abd05304 Pseudomon
c 17	66	9.4	1326	11	Abd05398 Pseudomon
18	65	9.3	77536	3	Aa14651 Nucleotid
c 19	64	9.2	9521	4	Aas07627 Streptoco
20	63.4	9.1	3354	8	AAL61181 Actinosyn

21	63.4	9.1	5862	10	AAD54224	Aad54224 Streptomy
22	63.4	9.1	52101	10	AAD54217	Aad54217 Streptomy
c 23	63.4	9.1	82746	8	AAL61224	Aal61224 Actinosyn
24	63.2	9.0	1227	10	ADC36139	Adc36139 Weed cont
25	63.2	9.0	1439	10	ADC36142	Adc36142 Weed cont
26	62.8	9.0	110000	4	AAI99682_36	Continuation (37 o
27	62.8	9.0	110000	4	AAI99683_36	Continuation (37 o
28	62.6	9.0	582	8	ACA23667	ACA23667 Prokaryot
29	62.4	8.9	819	8	ACA24088	ACA24088 Prokaryot
30	62.4	8.9	4770	8	AAL61172	Aal61172 Actinosyn
c 31	62.4	8.9	109519	5	AAS08693	Aas08693 Micromono
32	61.8	8.8	1134	12	ADJ35079	Adj35079 DNA encod
33	60.8	8.7	5877	6	ABs78701	ABs78701 S. citric
34	60.6	8.7	381	13	ACN62234	ACn62234 Cotton gy
35	60.2	8.6	23673	6	ABZ75344	Abz75344 Human R11
36	60.2	8.6	27541	4	AAD17185	Aad17185 Streptomy
37	60.2	8.6	125401	4	AAD17186	Aad17186 Streptomy
c 38	60	8.6	27541	4	AAD17185	Aad17185 Streptomy
c 39	60	8.6	125401	4	AAD17186	Aad17186 Streptomy
c 40	59.6	8.5	2721	2	AAQ10543	Aaq10543 BamHI J-I
c 41	59.6	8.5	2721	2	AAQ10212	Aaq10212 BamHI J-I
42	58.8	8.4	9975	8	AAL61173	Aal61173 Actinosyn
43	58.4	8.4	642	8	ABZ71163	Abz71163 S. muraya
c 44	58.4	8.4	36321	8	ABZ71131	Abz71131 Streptomy
45	58.2	8.3	1386	13	ADS56031	Ads56031 Bacterial

ALIGNMENTS

RESULT 1
AAL44719
ID AAL44719 standard; DNA; 699 BP.
XX
AC AAL44719;
XX
DT 03-MAY-2002 (first entry)
XX
DE S virginiiae butanolide binding protein coding sequence.
XX
KW Butanolide binding protein; gene expression induction; operator;
KW transgenic plant; antibiotic production; gene; ds.
XX
OS Streptomyces virginiiae.
FH Key Location/Qualifiers
FT CDS 1..699
FT /tag= a
FT /product= "butanolide binding protein"

WO200196581-A1.

20-DEC-2001.

15-JUN-2001; 2001WO-JP005096.

15-JUN-2000; 2000JP-00180466.

(KANF) KANEKA CORP.

Shimmo A, Kato K, Yamada Y, Nihira T, Shindo T;

WPI; 2002-098073/13.

P-PSDB; AAM48990.

Inducing expression of gene under regulation by operator at actinomycetes self-regulator provision site, useful in producing transformant tobacco for production of antibiotics e.g. virginiamycin.

Example 1; Page 46-48; 57pp; Japanese.

The present invention relates to a method of inducing the expression of a gene under the regulation of an operator at an actinomycetes self-

Db 2690 ACCACGATCGGCCGATCTTACGTGGGACGTTCCCGGGATACAGTGTGTCCTCCAG 2749
 Qy 514 GTCACTCCGACCGCCAGGACCTCGCCACCGGATCTCGGTGATGTGGAACACGCTGCTG 573
 Db 2750 ACGGTACGCGACTTACGAGACCTCGAACCGCTACGCGCTGCTGCAGAAACACATCCTG 2809
 Qy 574 CCCAGCATGTCGCGGCTCCATGCTGACCTGATCGAAACCGGAGAGCGGATCGGG 633
 Db 2810 CCGGCATCGCGTTCCTCCGCTGCTGCGCGCTGCTCTCCGAGGAGCGGAGCA 2869
 Qy 634 AAGGTCGCGCGG 646
 Db 2870 CGCCTCGCGCGG 2882

RESULT 3

ADE14794

ID ADE14794 standard; DNA; 4346 BP.

XX ADE14794;

XX ADE14794;

XX 29-JAN-2004 (first entry)

XX Streptomyces coelicolor partial genome sequence 2.

DE antibiotic-producing strain; antibiotic production; scbA gene; Scbr gene;

XX actinorhodin; undecylprodigiosin; gene; ds.

KW Streptomyces coelicolor.

XX Key

FH complement(1199..2143)

FT /tag= a

FT /product= "Streptomyces coelicolor Srbs protein"

FT /transl_except= (3149..3147, aa:Tyr)

FT 2261..2308

FT /tag= b

FT /product= "Streptomyces coelicolor Srbr protein"

FT complement(3024..3812)

FT /tag= c

FT /product= "Streptomyces coelicolor Srba protein"

FT /transl_except= (3150..3145, aa:Met)

FT US2003124644-A1.

XX 03-JUL-2003.

XX 23-OCT-2001; 2001US-00017471.

XX 23-OCT-2000; 2000US-0242561P.

XX (TAKA/) TAKANO E.

XX (BIBB/) BIBB M J.

XX Takano E, Bibb MJ;

XX WPI; 2003-810983/76.

XX P-PSDB; ADE14791, ADE14792, ADE14793.

XX Modifying an antibiotic-producing strain of Streptomyces coelicolor or

XX Streptomyces lividans to increase or alter the timing of antibiotic

XX production in the strain, comprises functionally deleting in the strain

XX the scbA or Scbr gene.

XX Claim 19; Fig 14; 33pp; English.

XX This invention relates to the novel modification of an antibiotic-

XX producing strain of Streptomyces coelicolor or Streptomyces lividans to

XX increase or to alter the timing of antibiotic production in the strain.

XX The method comprises functionally deleting in the strain the scbA or Scbr

XX gene. The method is useful in increasing and altering the timing of

XX antibiotic production (especially actinorhodin and undecylprodigiosin) in

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC Streptomyces species, particularly Streptomyces coelicolor or
 CC Streptomyces lividans. The present sequence is that of a region of the
 CC Streptomyces coelicolor genome, which encodes the ScbA, ScbB and ScbR
 CC proteins and which is related to the invention.
 XX
 SQ Sequence 4346 BP; 650 A; 1561 C; 1519 G; 616 T; 0 U; 0 Other;

Query Match 32.8%; Score 229; DB 10; Length 4346;
 Best Local Similarity 60.8%; Pred. No. 2.1e-30;
 Matches 373; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

Qy 34 CAGGAACGGCGCTCGCACGGGAGGATGTCGGGGGACCGCCCTCGGTCTTCGAC 93
 Db 2270 CAGGACCGGGCGATCCGACGGCGGAGATCTCTGGAGCGCGCGGAGGTTCTTCGAG 2329

Qy 94 GAGTACGGTTTCGAGCGCCACAGTGGCAGAGATCTCTCGGGGCTCGGTTCACCAAG 153
 Db 2330 AAGCAGGGCTACCAAGTGCACGATCAGGAGATCTCAAGTGGCCGGGTGACCAAG 2389

Qy 154 GCGCGATGATCTTCACATTCGTTCCAAAGGAAGAGCTGGCCCGCGGCTGCTGGCGGAG 213
 Db 2390 GGAGCGCTCTACTTCCACTTCCAGTCCAAAGGAAGAACTGGCGTGGGCGTCTTCGACGCC 2449

Qy 214 CAGACCTGACGTGGCGGTGCGGAATCCGGCTCCAAAGCGCGCAGGAACCTGTAGACCTC 273
 Db 2450 CAGGAACCAACAGCGCGCTTCGGAGCAACCCCTCCGGTGCAGAACTCATCGACATG 2509

Qy 274 ACCATGCTGGTCCCGCACCGCATGCTGCACGATCCGATCCTCGGGCGGGCACCGCGCTC 333
 Db 2510 GGCATGTTGTTCTGTACCGCTTGGCACGACGTCGTGGCGCGCGCGCGTGGCGCTC 2569

Qy 334 GCATGGACAGGGGCGGTGACATTCCTCCGACGCCAACCCGTTCCGGGAGTGGGGGAC 393
 Db 2570 TCATGGACAGCGCGGACGCGTCTCGATCGCGGAGACCCCTTCGCTCGTGGCACGAG 2629

Qy 394 ATCTGCGCCAGCTCTGCGGAGGACACAGGAACGGGGGAGGTGCTTCGACAGTGAAC 453
 Db 2630 ACATCTGAACTGCTGAAACCGAGCCAAAGGAAGACGGTGTGTCGCCCATGTGTC 2689

Qy 454 CGGAAAAAGACCGGGGACATTCATCGTGGCTGTGTTCCCGGGTCCAGGCGGTCTCCCGG 513
 Db 2690 ACCACGACTCGCGCGATCTCTACGTGGGACGTTTCGCGGGATACAGTTCGTGCCAG 2749

Qy 514 GTCACCTCCGACCGCGGACCTCGGCGCACCGGATCTCGGTGATGTGGAACACGCTGCTG 573
 Db 2750 ACGGTACGCGACTACCGAGACCTCGAAACCGCTACGCGCTGCTGAGAAGACATCTCTG 2809

Qy 574 CCCAGCATGTCGCGCGTCCATGCTGACCTGATCGAAACCGGCGGAGCGGATCGGG 633
 Db 2810 CCGGCATCGCGGTTCCCTCCGCTGTCGCGCGCTGATCTCTCCAGAGCGCGGAGCA 2869

Qy 634 AAGGTCGCGCGG 646
 Db 2870 CGCCTCGCGCGG 2882

RESULT 4
 AAZ99101/c
 ID AAZ99101 standard; DNA; 12905 BP.

XX AC AAZ99101;

XX 21-JUN-2000 (first entry)

XX S. fradiae tylosin biosynthetic pathway gene cluster antisense strand.

XX Gene cluster; macrolide antibiotic; tylosin; biosynthetic pathway;

XX recombinant; Streptomyces; hybrid antibiotic; ss.

XX Streptomyces fradiae.

XX Key

XX Location/Qualifiers

XX complement(294..968)

XX CDS

```

FT FT /tag= a
FT FT /label= ORF11
FT FT /gene= "tylQ"
FT FT /product= "regulatory protein"
FT FT 1315..3291
FT FT /tag= b
FT FT /label= ORF10
FT FT /gene= "tylP"
FT FT /product= "acyl-CoA dehydrogenase"
FT FT complement(3452..4069)
FT FT /tag= c
FT FT /label= ORF9
FT FT /gene= "tylO"
FT FT /product= "delta-TDP-glucose epimerase"
FT FT complement(4106..4873)
FT FT /tag= d
FT FT /label= ORF8
FT FT /gene= "tylF"
FT FT /product= "macrosin-O-methyltransferase"
FT FT 5137..6399
FT FT /tag= e
FT FT /label= ORF7
FT FT /gene= "tylH1"
FT FT /product= "cytochrome P450"
FT FT 6435..6680
FT FT /tag= f
FT FT /label= ORF6
FT FT /gene= "tylH2"
FT FT /product= "ferrodoxin"
FT FT /note= "the termination codon of this ORF overlaps with
FT FT the initiation codon of ORF5"
FT FT 6677..7687
FT FT /tag= g
FT FT /label= ORF5
FT FT /gene= "tylD"
FT FT /product= "delta-TDP-glucose 4-6 dehydratase"
FT FT /note= "the initiation codon of this ORF overlaps with
FT FT the termination codon of ORF6"
FT FT 7705..8892
FT FT /tag= h
FT FT /label= ORF4
FT FT /gene= "tylE"
FT FT /product= "methyltransferase"
FT FT 8938..10206
FT FT /tag= i
FT FT /label= ORF3
FT FT /gene= "tylN"
FT FT /product= "glycosyl transferase"
FT FT complement(10276..11118)
FT FT /tag= j
FT FT /label= ORF2
FT FT /gene= "tlrB"
FT FT /product= "tylosin resistance protein"
FT FT 11475..12671
FT FT /tag= k
FT FT /label= ORF1
FT FT /gene= "ddcA"
FT FT /product= "D-alanine carboxypeptidase"
FT FT
XX PN WO200008168-A1.
XX
XX PD 17-FEB-2000.
XX
XX PF 02-AUG-1999; 99WO-ES000248.
XX
XX PR 05-AUG-1998; 98ES-00001682.
XX
XX PA (ANTI ) ANTIBIOTICOS SAU.
XX
XX PI Fouces Martinez R, Mellado Duran E, Diez Garcia B;
XX PI Esteban Morales M, Bernasconi E, Barredo Fuente JL;
XX DR WPI; 2000-205718/18.

```

```

DR P-PSDB; AAY83784, AAY83785, AAY83786, AAY83787, AAY83788, AAY83789,
DR AAY83790, AAY83791, AAY83792, AAY83793, AAY83794.
XX
PT Increasing tylosin and/or hybrid antibiotic production in Streptomyces
PT species by introducing genes of the Streptomyces fradiae tylosin
PT biosynthetic cluster.
XX
PS Claim 7; Page 33-43; 63pp; Spanish.
XX
CC This sequence corresponds to the antisense strand for the gene cluster
CC encoding proteins involved in the macrolide antibiotic tylosin
CC biosynthetic pathway. The invention relates to methods of increasing the
CC level of production of tylosin and/or hybrid antibiotics in Streptomyces
CC spp. comprising culturing cells that have been transformed with a vector
CC containing a Streptomyces fradiae genomic DNA fragment. The method is
CC useful either for producing recombinant S. fradiae strains that
CC overproduce the antibiotic tylosin or for producing recombinant strains
CC of other Streptomyces spp. that will produce novel hybrid antibiotics
XX
SQ Sequence 12905 BP; 1859 A; 4723 C; 4547 G; 1776 T; 0 U; 0 Other;
Query Match 18.0%; Score 126; DB 3; Length 12905;
Best Local Similarity 56.2%; Pred. No. 7.4e-13;
Matches 342; Conservative 0; Mismatches 250; Indels 16; Gaps 5;
QY 28 GTGCGACAGGAACGGGCGGTCCGACGGGAGGCGATCGTCGGGCGACCGCTCGGTC 87
DB 965 GCGCGACAGGAACGGGCGGCGCCACACCGCGCGGACGATCGT---CGGCGCGCGCGG 909
QY 88 TTCGACGAGTACGGTTGAGGCGCCACAGTGGCAGAGATCCTCTCGGCGGCGCTCGGTC 147
DB 908 TTCGACGAACTGGGCTACGAGCGGACCACTATCGGGAGATCTCTCAAGAGTCCGGGTG 849
QY 148 ACCAAGGCGCGATGTACTTCCACTTCCAGGAGAGAGTGGCCCGCGCGGTCG 207
DB 848 ACCAAGGAGCGCTGTACTTCCACTTACGTCCAGGAGCAGCTCGCGGAGAGTGGTG 789
QY 208 GCGGAGCAGACCTCGACAGTG---GCGGTGCGCGAATCCGGCTCCAAAGCGCAGGAAC 264
DB 788 ACGAGTCAGTCCGGGCGGTGCGCGGTGGAGGAACAGCGACTCGTCTCCAGCAGATC 729
QY 265 GTAGACCTCACCATGCTGTGTCGCCA---CGGCATGCTGCAGATCCGATCTCTCGGGCG 321
DB 728 ATCGACGAGACGCTGTCTGTGTCGCAACTGCTCAGCAGGGCGATCCGCTGGTGGCGGC 669
QY 322 GGCACGCGGCTCGCACTGGACCGAGGGGC-----GGTGGACTTCTCCGACGCCAACCCG 375
DB 668 AGTGTCGGCTCAGCGTGAGCGCGGCGCGCCCGGAGGACGGCTGGACCGCAGGGCGCG 609
QY 376 TTCGGCGAGTGGGGGACATCTGCGCCAGCTCTTGGCGGAGGACAGGAACGGGGGAG 435
DB 608 ATGCAAGAGTGGATCGGCCACGCGAGGATCTCTCAGAAGGGCGAGGCGGCGGTAA 549
QY 436 GTGCTTCGCGACGTGAACCCGAAAGACCGCGGACTTCATGTCGGGCTGCTTACCGGG 495
DB 548 CTGCTGCGCGGCTCGATGTCGACGCGGTGGCGAGAAATGCTGTGGCGGCTTTCACCGGT 489
QY 496 CTCGAGGCGGTCTCCGGGTACCTCCGACCGCAGGACCTCGGCCACCGGATCTCGGTG 555
DB 488 GCCCAGATCTGTGAAACATCTGACCGGGCAGCGCGATCTCTGTGAGCGGGTGACCGAC 429
QY 556 ATGTGAACCACTGCTGCCAGCATCGTGCCGGGTCCATGCTGACTGAGCTGGAATCGAAACC 615
DB 428 ATGCACCGGCATCTCATGACCTCGGTC-CGGTGCCGCGGTGCTGTGTGCGGCTCGACTTC 370
QY 616 GCGGAGGA 623
DB 369 TCCGCGGA 362

```

RESULT 5
AAZ99100
ID AAZ99100 standard; DNA; 12905 BP.

```

XX AC AAZ99100;
XX DT 21-JUN-2000 (first entry)
XX DE S. fradiae tylosin biosynthetic pathway gene cluster sense strand.
XX DT
XX DE Gene cluster; macrolide antibiotic; tylosin; biosynthetic pathway;
XX DT recombinant; Streptomyces; hybrid antibiotic; ss.
XX DE Streptomyces fradiae.
XX DT
XX DE Location/Qualifiers
XX DT complement(235..1431)
XX DT /tag= a
XX DT /label= ORF1
XX DT /gene= "ddca"
XX DT /product= "D-alanine carboxypeptidase"
XX DT 1788..2630
XX DT /tag= b
XX DT /label= ORF2
XX DT /gene= "tlrB"
XX DT complement(2700..3968)
XX DT /tag= c
XX DT /label= ORF3
XX DT /gene= "tylN"
XX DT /product= "glycosyl transferase"
XX DT complement(4014..5201)
XX DT /tag= d
XX DT /label= ORF4
XX DT /gene= "cyle"
XX DT /product= "methyltransferase"
XX DT complement(5219..6229)
XX DT /tag= e
XX DT /label= ORF5
XX DT /gene= "tylD"
XX DT /product= "delta-TDP-glucose 4-6 dehydratase"
XX DT /note= "the initiation codon of this ORF overlaps with
XX DT the termination codon of ORF6"
XX DT complement(6226..6471)
XX DT /tag= f
XX DT /label= ORF6
XX DT /gene= "tylH2"
XX DT /product= "ferrodoxin"
XX DT /note= "the termination codon of this ORF overlaps with
XX DT the initiation codon of ORF5"
XX DT 6507..7769
XX DT /tag= g
XX DT /label= ORF7
XX DT /gene= "tylH1"
XX DT /product= "cytochrome P450"
XX DT 8033..8800
XX DT /tag= h
XX DT /label= ORF8
XX DT /gene= "cylF"
XX DT /product= "macrosin-O-methyltransferase"
XX DT 8837..9454
XX DT /tag= i
XX DT /label= ORF9
XX DT /gene= "tylO"
XX DT /product= "delta-TDP-glucose epimerase"
XX DT complement(9615..11591)
XX DT /tag= j
XX DT /label= ORF10
XX DT /gene= "tylP"
XX DT /product= "acyl-CoA dehydrogenase"
XX DT 11938..12612
XX DT /tag= k
XX DT /label= ORF11
XX DT /gene= "tylQ"
XX DT /product= "regulatory protein"

```

```

PN WO200008168-A1.
XX
XX 17-FEB-2000.
XX
XX 02-AUG-1999; 99WO-ES000248.
XX
XX 05-AUG-1998; 98ES-00001682.
XX
XX (ANTI ) ANTIBIOTICOS SAU.
XX
XX Fouces Martinez R, Mellado Duran E, Diez Garcia B;
XX Esteban Morales M, Bernasconi E, Barredo Fuente JL;
XX
XX WPI; 2000-205718/18.
XX
XX P-PSDB; AAY83784, AAY83785, AAY83786, AAY83787, AAY83788, AAY83789,
XX AAY83790, AAY83791, AAY83792, AAY83793, AAY83794.
XX
XX Increasing tylosin and/or hybrid antibiotic production in Streptomyces
XX species by introducing genes of the Streptomyces fradiae tylosin
XX biosynthetic cluster.
XX
XX Claim 7; Page 26-32; 63pp; Spanish.
XX
XX This sequence corresponds to the sense strand of the gene cluster
XX encoding proteins involved in the macrolide antibiotic tylosin
XX biosynthetic pathway. The invention relates to methods of increasing the
XX level of production of tylosin and/or hybrid antibiotics in Streptomyces
XX spp. comprising culturing cells that have been transformed with a vector
XX containing a Streptomyces fradiae genomic DNA fragment. The method is
XX useful either for producing recombinant S. fradiae strains that
XX overproduce the antibiotic tylosin or for producing recombinant strains
XX of other Streptomyces spp. that will produce novel hybrid antibiotics
XX
XX Sequence 12905 BP; 1776 A; 4548 C; 4722 G; 1859 T; 0 U; 0 Other;
XX
XX Query Match 18.0%; Score 126; DB 3; Length 12905;
XX Best Local Similarity 56.2%; Pred. No. 7.4e-13;
XX Matches 342; Conservative 0; Mismatches 250; Indels 16; Gaps 5;
XX
XX QY 28 GTGCGACAGAGAACGGGCGCTCGCACGCGGAGCGGCTGTCGGGAGCGGCTCGGTC 87
XX DB 11941 GCGCGACAGAGAACGGGCGCTCGCACGCGGAGCGGCTGTCGGGAGCGGCTG 11997
XX
XX QY 88 TTGCGACAGTACGGGTTGAGGCGCGCCACAGTGGGAGAGATCTCTTCGGGGGCTCGGTC 147
XX DB 11998 TTGCGACAGTACGGGTTGAGGCGCGCCACAGTGGGAGAGATCTCTTCGGGGGCTCGGTC 12057
XX
XX QY 148 ACCAAGGGCGGATGTACTTCCACTTCGCTTCCAGGAGAGAGTGGCCCGCGGCTGCTG 207
XX DB 12058 ACCAAGGGCGGATGTACTTCCACTTCGCTTCCAGGAGAGAGTGGCCCGCGGCTGCTG 12117
XX
XX QY 208 GCGGAGCAGACCTTGACCGTG---GCGGTGCGCGGATCCGGGTCCAGGCGCAGGAATCG 264
XX DB 12118 ACGAGTCAGCTCGGGCGCTGTCGCCCGGTGGAGAACAGCGACTCTCTCCAGCAGATC 12177
XX
XX QY 265 GTAGACCTCACCATGCTGTGTCGCCA---CGGATGCTGCAGATCCGATCTTCGGGGCG 321
XX DB 12178 ATCGACGAGACGCTGCTGTGCGCAACTGCTCAGCAAGGGGATCCGCTGTGTGCGCGG 12237
XX
XX QY 322 GGCACGCGGCTCGCACTGCACACAGGGGC-----GGTGGACTTCTCCGACGCCACCCG 375
XX DB 12238 AGTGTCCGGCTCAGGTGGAGCGCGGCGCCCGCCAGGAGCGGCTTGACCGAGGGCGCGG 12297
XX
XX QY 376 TTGCGGAGTGGGGACATCTGCGCCAGCTCTGCGGGAGGACACAGGAACGGGGGGAG 435
XX DB 12298 ATGCAAGAGTGGATCGGCACACGCGAGGATCTCTCAGAAAGGCGGCGGGGTGAA 12357
XX
XX QY 436 GTGCTTCGCACGTAACCCGAAAGACCGCGACTTCATCGTGGGCTGCTTCACCGGG 495
XX DB 12358 CTGCTGCCGCGCTCGATGTGACGCGCGGTGGGAGAAATCTGTGCGCGGTTTACCGGT 12417
XX
XX QY 496 CTCAGGCGGCTCTCCGGGTTCACCTCCGACCCCGAGGACCTCGGCCACCGATCTCGGTG 555

```

Db 12418 GCCAGATCTCTGCAACATCTGACGGGCGACGCGGATCTGCTGGAGCGGTGACCGAC 12477
QY 556 ATGTGGAACACGTCGTCGCCAGCATCGTCGGGTCCTCATCTGACCTGATCGAAACC 615
Db 12478 ATGCACCGCATCTCATGACCTCGGTC-CGGTCCCGCGCTGCTGTCGGCTCGACTTC 12536
QY 616 GCGGAGGA 623
Db 12537 TCCGCGGA 12544

RESULT 6

ABV75877/c
ID ABV75877 standard; DNA; 1819 BP.
XX
AC ABV75877;
XX
DT 07-FEB-2003 (first entry)
XX
DE Streptomycetes coelicolor type II thioesterase scot gene.
XX
KW Thioesterase; enzyme; Scot; polyketide; antibiotic; gene; ds.
XX

Streptomycetes coelicolor.

Key Location/Qualifiers
CDS 748..1554
FT /tag= a
FT /product= "Type II thioesterase"
FT /note= "a polynucleotide comprising the CDS is
FT specifically claimed in Claim 10"

WO200279461-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-FL000027.

28-MAR-2001; 2001PL-00346786.

(KUCZ/) KUCZEK K.
(KOTO/) KOTOWSKA M.
(PAWL/) PAWLICK K.
(WIERN/) WIERNIK D.

Kuczek K, Kotowska M, Pawlik K, Wiernik D;

WPI; 2003-058432/05.

P-PSDB; ABP55143.

New thioesterase type II (TE II) protein, useful for expression of a protein having TE II activity or for increasing biosynthetic efficiency of a polyketide synthase complex.

Claim 1; Page 11-12; 13pp; English.

The present sequence is that of the type II thioesterase (TE II) scot gene of Streptomycetes coelicolor A3(2) M145. The gene was isolated from a genomic library using a probe based on the central region of the acyltransferase domain from type I polyketide synthase, including the active site region. DNA reacting with the probe, obtained as 2 restriction fragments, was cloned and sequenced. Correction of frame-shift mistakes provided the present sequence, which includes an open reading frame (also claimed), designated scot. This encodes a 268-amino acid protein (see ABP55143) having a mol.wt. of 28,686 and a pI of 6.17, which has sequence similarity to known TE II enzymes. The isolated gene also includes a region active in transcriptional regulation of scot. The scot enzyme is active in polyketide, especially macrolide, biosynthesis when associated with a multienzyme complex of a polyketide synthase. Its relatively broad substrate tolerance will facilitate the production of novel, heterologous polyketide synthase complexes. The scot gene may be transferred to a different streptomycete to increase the efficiency of production of a polyketide (possibly an antibiotic)

XX SQ Sequence 1819 BP; 246 A; 683 C; 620 G; 270 T; 0 U; 0 Other;
Query Match 12.7%; Score 88.6; DB 8; Length 1819;
Best Local Similarity 54.2%; Pred. No. 2.1e-06;
Matches 215; Conservative 0; Mismatches 164; Indels 18; Gaps 1;
QY 34 CAGGAACCGGCGGTCGCGACGCGGCGGATGTCGGGGGAGCGGCTCGGTCTTCGAC 93
Db 500 CAGGAGCGGCGGCGACGTCAGCGTCTGATTTCTCTCGCGCGCGAGGTGTTGAC 441
QY 94 GAGTACGGGTTGAGGCGCGCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTACCCAG 153
Db 440 CAGGAGGGGTTGCGGCGCGCGTCTGCTCAGATGATCAGTCCCGGCGGGGTGAGCAAC 381
QY 154 GCGCGATGTACTTCCACTTCGCTTCCAAAGGAAGAGCTGGCCCGGGGCTGTGCGCGGA- 212
Db 380 GGGGCACTGCACTTCCACTTCGCGACAAAGAAACGCGGTGGCGAGCGGTGAGGGCGAG 321
QY 213 -----GCAGACCTGCACGCTGGCGGTGCGCGGAATCGGGCTTCAAGGG 255
Db 320 GCGTTGTCGGTGTCTCCGGCAGATGCCACGCGTGGCGGAGGGGGCCACGCTCGCTC 261
QY 256 CAGGAACCTGTAGACCTCACCATGCTGTGTCGCCACGGCATGCTGCACGATCGGATCTCTG 315
Db 260 CAGTCCCTCGTCGACACCTCGCACGCTGGCGCCACGACATCCAGGACGATGTCGTCTG 201
QY 316 CGGGCGGCGACGCGGCTCGCACTGACACGAGGGGGCGGTGGAGCTTCTCGACGCCAACCGG 375
Db 200 CCGGCGGCTTTCGGGCTGAGCGGGGACACGACCTGGAAGGACGCGCGGATGTCGCCGC 141
QY 376 TTCGGCGAGTGGGCGCACATCTGCCCGCCAGTCTCTGG 412
Db 140 CACTGGGTGGACTGGGTGAGCAGCGGCTGACCGTGG 104

RESULT 7

AAS07627

ID AAS07627 standard; DNA; 9521 BP.

XX AC AAS07627;

XX 11-SEP-2003 (revised)

DT 23-OCT-2001 (first entry)

XX Streptococcus coelicolor genomic DNA from plasmid SCP1.

XX SCP1; ds; methylenomycin cluster; mmc; MmyR; MmfP; MmfH; MmfL; MmfR;

KW MmyT; MmyO; MmyG; MmyU; Mmr; heterologous gene expression.

XX Streptomycetes coelicolor.

Key Location/Qualifiers

CDS complement(796..1407)

FT /tag= a

FT /product= "MmyR"

FT /note= "This sequence is specifically claimed"

FT /partial

FT /note= "The mmyR polypeptide may be encoded from the ATG at 1387-1399"

FT complement(1390..1557)

FT /tag= b

FT /label= mmyR_promoter

FT /note= "This sequence is specifically claimed, and

FT optionally excludes nucleotides 1390-1409"

FT complement(1558..2355)

FT /tag= c

FT /product= "MmfP"

FT /note= "This sequence is specifically claimed"

FT /partial

FT /note= "The mmfP polypeptide may be encoded from the ATG at 2350-2352"

FT complement(2352..3554)

FT CDS

```
FT FT /*tag= d
FT FT /product= "MmFH"
FT FT /note= "This sequence is specifically claimed"
FT FT complement(3551..4612)
FT FT /*tag= e
FT FT /product= "MmFL"
FT FT /note= "This sequence is specifically claimed"
FT FT 4613..4806
FT FT /*tag= f
FT FT /label= "MmFR_promoter"
FT FT /note= "This sequence is specifically claimed"
FT FT complement(4613..4806)
FT FT /*tag= g
FT FT /label= "mmfL/mmFH_promoter"
FT FT /note= "This sequence is specifically claimed"
FT FT 4807..5451
FT FT /*tag= h
FT FT /product= "MmFR"
FT FT /note= "This sequence is specifically claimed"
FT FT 5452..5675
FT FT /*tag= i
FT FT /label= "mmyTOG_promoter"
FT FT /note= "This sequence is specifically claimed"
FT FT 5676..6401
FT FT /*tag= j
FT FT /product= "MmyT"
FT FT /note= "This sequence is specifically claimed"
FT FT 6432..7553
FT FT /*tag= k
FT FT /product= "MmyO"
FT FT /note= "This sequence is specifically claimed"
FT FT 7536..8817
FT FT /*tag= l
FT FT /product= "MmyG"
FT FT /note= "This sequence is specifically claimed"
FT FT complement(8780..9151)
FT FT /*tag= m
FT FT /product= "MmyJ"
FT FT /partial
FT FT /note= "The mmyJ polypeptide may be encoded from the ATG
FT FT at 9113-9115"
FT FT 9333..9521
FT FT /*tag= n
FT FT /product= "Mmr"
FT FT /partial
FT FT /note= "No stop codon"
FT FT
XX WO200148228-A1.
XX
XX 05-JUL-2001.
XX
XX 20-DEC-2000; 2000WO-GB004972.
XX
XX 23-DEC-1999; 99GB-00030477.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Chater KF, Bruton CJ, O'rouke SJ, Wietzorrek AW;
XX
XX WPI: 2001-425675/45.
XX
XX P-PSDB; AAU04039, AAU04040, AAU04041, AAU04042, AAU04043, AAU04044,
XX AAU04045, AAU04046.
XX
XX Novel expression cassette for expressing a nucleic acid of interest,
XX derived from the regulatory region of methylenomycin gene cluster of SCP1
XX plasmid of Streptomyces coelicolor A3(2).
XX
XX Example 3; Fig 7; 142pp; English.
XX
XX The sequence represents an expression cassette (present on plasmid SCP1)
XX and is the regulatory region of the methylenomycin cluster (mnc) from
XX Streptomyces coelicolor A3(2), which encodes the MmyR, MmFp, MmFL,
XX MmFR, MmyT, MmyO, MmyG, MmyU and partial Mmr polypeptides. The expression
```

```
CC cassette is useful for expressing a nucleic acid of interest,
CC substantially only when the host cell culture reaches high cell density
CC at or close to the stationary phase of host cell culture. In particular
CC the system is useful in regulating methylenomycin production. Reduced or
CC no expression of the nucleic acid of interest is observed earlier in
CC growth, avoiding toxic effects of some gene products on growth and the
CC system does not require addition of exogenous inducer. The methylenomycin
CC cluster naturally present on a highly transmissible plasmid permits
CC properly regulated expression in diverse Streptomyces host and the
CC expression is driven by a strong promoter, leading to high yield of the
CC desired end product. (Updated on 11-SEP-2003 to standardise OS field)
XX
XX Sequence 9521 BP; 1399 A; 3223 C; 3416 G; 1483 T; 0 U; 0 Other;
```

```
Query Match 12.3%; Score 85.8; DB 4; Length 9521;
Best Local Similarity 50.1%; Pred. No. 5.5e-06;
Matches 281; Conservative 0; Mismatches 262; Indels 18; Gaps 2;

QY 30 GCACACAGGAACGGGCGGTCCGCACGGCGGAGGCGATCGTGGGGGACGCCCTCGGTCTTT 89
DB 4872 GCAGCAGGAGCGGTCTGATCAAGACCGGGGCCAGATCTTGGAGGCGCGTTCGAGATCTT 4931

QY 90 CGACGAGTACGGGTTCCGAGCGCCACAGTGGCAGAGATCCTCTCCGGGCGCTCGGTTCAC 149
DB 4932 CGGTCCGCGGCTACCGAGGGGCTTCCGTCAAGGACGTTGCCGAGCGTTCGGCATGAC 4991

QY 150 CAAGGGCGCGATGTACTTCCACTTTCCTTCCAAAGGAAGAGCTGGCGCGCGGCTGTGGC 209
DB 4992 CAAGGGCGGTTACTTCCACTTCCCGCAGCAAGGAATCCTTGGCCATCGCGTGGTGA 5051

QY 210 CGACGAGACCTTGACGTGGCGG-----TGCCGAATCCGGCTCAAGGC----- 254
DB 5052 GGAGCACTACGGCGGCTGGCGCGCAGCATGGAAGAGATCCGCATCAGGGCTTCACACC 5111

QY 255 ----GCAGGAACCTGGTAGACCTCACCATGCTGGTCCGCCACGCGATCTGCACGATCCGAT 311
DB 5112 GCTGGAGACGTCGAGGAGATGCTCATCGCGGGCGCAGGCTTCCGGACACGCCCGT 5171

QY 312 CCTGGGGGGGCGACGCGCTCGCATCTGGACACGAGGGGGGGTGGACTTCTCCGACGCCAA 371
DB 5172 GATGAGGCGGCTGGCGCGCTGCAGAGTACGCGCGCTTTCATCGACGCGAGCTGCCCT 5231

QY 372 CCGGTTGGCGAGTGGGGGACATCTGCCCGCAGCTCTTGGGGGAGGACAGAGAAACGGGG 431
DB 5232 GCCCTACGTGGACTGGACCCACCTGCTGGAGGTGCGTTCGAGGAGCGCCCGTGGAGCCGG 5291

QY 432 GGAGGTGCTTCGCGACGTGAACCCGAAAGACCGCGCATTCATCTCGGCTGCTTCAC 491
DB 5292 CCAGTTGGGGGGGCGTTCGATCCCGCAGCATGCCCCGTTCCCTGGTGGCGGCTTCTT 5351

QY 492 CGGGCTCCAGGCGGCTCTCCCGGCTCACCTCCGACCGCCAGGACCTCGGCCACCGGATCTC 551
DB 5352 CGGCATGCAGCAGCTCTCCGACAACTGCACACGCGGCGGACATCATGGAGCGGTGGCA 5411

QY 552 GGTGATGTGGAACACGAGTGTCT 572
DB 5412 GGAGCTGGCGGAGCTGATGTT 5432
```

```
RESULT 8
AAV68520
ID AAV68520 standard; DNA; 4257 BP.
XX
XX AAV68520;
XX AC
XX 29-JAN-1999 (first entry)
XX
XX The nucleotide sequence of the Herpes simplex virus ICP4.
XX
XX Herpes simplex virus ICP4; HSV ICP4; modulator; apoptosis; stimulation;
XX inhibition; HSV infection; ss.
XX
XX Herpes simplex virus unknown type.
OS
```


Db 1927 ATCTTGGCGCGCTGGGGCGGTGTCTCCGCGCGCCCGCTCCCGCGGGGGCGACGAC 1986
QY 163 TACTTCCACTTGCCTTCCAAAGAGAGCTGGCCCGCGGTGTCTGGCCGAGCAGACCCCTG 222
Db 1987 CCGAGCGCGCGCCGACCGCGAGCGCGAGCAGCGCGGGCGCGCGCCGAGCGCGC 2046
QY 223 CAGCTGGCGGTGCGGAAATCCGGCTCAAGGGCGAGAACTGTGTAGACTTCACTATGCTG 282
Db 2047 CGGTGGCGGTGAGTGCCTTGGCGCTGCGCGGATCTGTGAGGCGTGTGGCCGAGGCG 2106
QY 283 GTGCCCGAGGATGCTGCAAGATCGATCTGCGGGCGGGGACGCGGTCTGCACTGGAC 342
Db 2107 TTGACGGCGACCTGGCGCGGTTCGGGGCTGGCGGGCGCGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGGCGGTGACTTCTCCGACGCAACCCCGTTTCGGCGAGTGGGGCGACATCTCGGC 402
Db 2167 CGGCGGAGGACCGCGGGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTGGCGAGGACAGGAAACGGGGGAGGTCTTCGCACTGTGAACCGGAAAG 462
Db 2227 CTGGCGCGGTGCTGGCGAGTGCCTGCTGCGCGAGCGCTGTGCTCATGCGCGCTG 2286
QY 463 ACGGCGACTTATCTGCTGCTGCTTACGCGGCTCCAGCGGTCTCCGGGTCACTCC 522
Db 2287 CGGGGGGACTGCGCTGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 2346
QY 523 GACCGCGAGACTCTCGGCGACCGGATCTCGGTGATGTGAACACCGTGTGCGCGCATC 582
Db 2347 AGCTGTGCGCGGGCGCTTGGCGCGCGCTGCGCGGGACCGCGCTTGGCGACTC 2406
QY 583 GTGCCGCTTCATGCTGATGATGAAACCGGCGAGAGCGGATCGGGAAGTTCGCG 642
Db 2407 GCGGCGCGCGCGCGCGGACCTGCTGTTGAACACGAGAGCTGCGCGCGCTGCTGG 2466
QY 643 GCGGCGCGCGCGCGCGGCTGCGGAGCTCGGAGCTTCCGAGCGCGCTCCGAC 693
Db 2467 GCGGCGCGCGCGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2517

RESULT 10
ACA37971 standard; DNA; 1224 BP.
XX ID ACA37971 standard; DNA; 1224 BP.
XX AC ACA37971;
XX CT ACA37971;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #19628.
XX KS Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX LX Mycobacterium avium.
XX OS W0200277183-A2.
XX PN 03-OCT-2002.
XX PP 21-MAR-2002; 2002W0-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362899P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX PI WPI; 2003-029926/02.

DR P-PSDB; ABU34101.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 25841; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1224 BP; 138 A; 451 C; 485 G; 150 T; 0 U; 0 Other;

Query Match 9.8%; Score 68.2; DB 8; Length 1224;
Best Local Similarity 46.4%; Pred. No. 0.0065;
Matches 306; Conservative 0; Mismatches 338; Indels 15; Gaps 2;
QY 36 GGAACGGGCGCTCGCACCGCGGAGCGGATCTGCGGGGAGCGCGCTCTTCGACGA 95
Db 522 GCAGCGGGTGTGTCTCGCGGTGCGCGATTTGCGGCTTCCCGCGCGCGAGGTGTGCGCGT 581
QY 96 GTACGGGTTTCGAGGGCGCCACAGTGGCAGAGATCTCTTCGCGGGCTTCGAGGAGG 155
Db 582 CCGGTGCGCGAGGCGCATCACCTCGTGACCGCGCGGGGTGCGCGCGCGCGCGTGC 641
QY 156 CGCGATGTACTTCCACTTCCGCTTCCAGGAAGAGCTGGCGCGCGGGGTGCGCGCGAGCA 215
Db 642 GCACCTTCGGAATTCACCGACGCGCGGGGTGTGATGGTTCGCGGGGTTCGCGGGCGC 701
QY 216 GACCTTCGACGTGGCGGTGCGGAATTCGCGCTCCAAAGGCGCGAGAACTTGTAGACCTTAC 275
Db 702 GCGCTCTGTAACCGCGCGGTGTCCGCGCGCGCGCGAGCTGCGCGCGCGCGCGGTGC 761
QY 276 CATGCTGTGTCGCGCCACGGCATGCTGCACGATCCGATCTTCGCGCGCGGGACGCGGTTCG 335
Db 762 CTGTGCTGCACGCGCAGCGCGCCCAAAAAACAGCTTCGACCTTGGCGGAAACCGCAGCGCGCG 821
QY 336 ACTGGAACGAGGGCGGTGGAATCTTCCGAC-----GCCAACCGTTCGCGCA 383
Db 822 CCGCGCTTACGTGGCGGTGCGCTTACCTGGAACGGAATGAGACCTTGGCGCTACGCGCGCGCG 881
QY 384 GTGGGCGGACATCTGCGCGCGCGCTCTTGGCGGAGGACAGGAAACGGGGGAGGTGCTTCC 443
Db 882 CCTGTTGATCTGCGCGTCCCGGGCGGATGACGCTGCGCGGAGGTGCTGCGCGGTGCGGTTGCC 941

```
QY 444 GCACGTGAACCCGAAAGACCGCGACTTCATCGTCGGTCTTCCACCGGCTCCAGGC 503
Db 942 GGCCATCTACGTCCGCTCCGATCGGACCGGGAACAGGGCTCAACGGCTGCCGT 1001
QY 504 GGTCTCCCGGGTCACTCCGACCGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGAA 563
Db 1002 GGTCAACGCGGGCGGCGCATGTGTGTCGCCGACGCGCACTTGACGCGGAGTTGTGCG 1061
QY 564 CCACGTGCTGCCAGCATGTGCGCG---CGTCCATGCTGACCTGATCGAATCGAAGCGGCA 620
Db 1062 CCGGGAGGTGGCGGCGCTGTGGCGACCCGCGGCTGCGGCGGATACGACGCGCCG 1121
QY 621 GGAGCGGATCGGAAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCG 679
Db 1122 CCGCGGGTTCGGGATCCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATCG 1180

RESULT 11
AD051695/c
ID AD051695 standard; DNA; 32329 BP.
XX AC AD051695;
XX DT
XX 15-JUL-2004 (first entry)
XX DE Streptomyces cattleya NRRL 8057 thienamycin biosynthetic gene cluster.
XX KW Beta-lactam synthetase; carbapenem antibiotic; thienamycin;
KW Gram-positive pathogenic bacteria; Gram-negative pathogenic bacteria;
KW infectious disease; therapy; gene; ds.
XX OS Streptomyces cattleya.
XX FH
XX Key Location/Qualifiers
FT CDS complement(2..751)
FT FT /*tag= a
FT FT /product= "Streptomyces cattleya orfX protein"
FT FT /partial
FT FT /note= "No stop codon"
FT CDS 1034..1924
FT FT /*tag= b
FT FT /product= "Streptomyces cattleya orfX1 protein"
FT FT complement(1940..2719)
FT CDS /*tag= c
FT FT /product= "Streptomyces cattleya orfX2 protein"
FT FT complement(2850..3932)
FT FT /*tag= d
FT FT /product= "Streptomyces cattleya orfZ1 protein"
FT FT 4117..4755
FT CDS /*tag= e
FT FT /product= "Streptomyces cattleya orfZ2 protein"
FT FT 4945..4936
FT FT /*tag= f
FT FT /product= "Streptomyces cattleya orfZ3 protein"
FT FT /partial
FT FT /note= "No start codon"
FT CDS 6288..7172
FT FT /*tag= g
FT FT /product= "Streptomyces cattleya orf1 protein"
FT FT complement(7156..8139)
FT CDS /*tag= h
FT FT /product= "Streptomyces cattleya orf1b protein"
FT FT /partial
FT FT /note= "No start codon"
FT CDS complement(8136..8927)
FT FT /*tag= i
FT FT /product= "Streptomyces cattleya orf1c protein"
FT FT 9171..9845
FT CDS /*tag= j
FT FT /product= "Streptomyces cattleya orfA protein"
FT FT complement(9767..11197)
FT CDS /*tag= k

/product= "Streptomyces cattleya orf2 protein"
/complement(11289..12740)
/*tag= l
/product= "Streptomyces cattleya orf2c protein"
/partial
/note= "No start codon"
complement(12737..14782)
/*tag= m
/product= "Streptomyces cattleya orf3 protein"
complement(14838..16262)
/*tag= n
/product= "Streptomyces cattleya orf4 protein"
complement(16234..17610)
/*tag= o
/product= "Streptomyces cattleya orf5 protein"
complement(17612..18715)
/*tag= p
/product= "Streptomyces cattleya orf6 protein"
complement(18754..20172)
/*tag= q
/product= "Streptomyces cattleya orf7 protein"
/partial
/note= "No start codon"
complement(20169..21623)
/*tag= r
/product= "Streptomyces cattleya orf8 protein"
22038..22817
/*tag= s
/product= "Streptomyces cattleya orf8d protein"
complement(22912..23634)
/*tag= t
/product= "Streptomyces cattleya orf9 protein"
23744..24733
/*tag= u
/product= "Streptomyces cattleya orf10 protein"
24784..25983
/*tag= v
/product= "Streptomyces cattleya orf11 protein"
/partial
/note= "No start codon"
26146..26952
/*tag= w
/product= "Streptomyces cattleya orf12 protein"
complement(26980..27393)
/*tag= x
/product= "Streptomyces cattleya orf13 protein"
complement(27614..28156)
/*tag= y
/product= "Streptomyces cattleya orfY protein"
/partial
/note= "No start codon"
28224..28412
/*tag= z
/product= "Streptomyces cattleya orfW1 protein"
28485..29246
/*tag= aa
/product= "Streptomyces cattleya orfW2 protein"
29382..29567
/*tag= ab
/product= "Streptomyces cattleya orfW3 protein"
29766..30197
/*tag= ac
/product= "Streptomyces cattleya orfW4 protein"
30695..32329
/*tag= ad
/product= "Streptomyces cattleya orfW5 protein"
/partial
/note= "No stop codon"

US2004038250-A1.
26-FEB-2004.
XX
```


Db 4806 GCGTCGCGCGCGCGCACCGCGCGAGCGCGCGGTGCGCGCGGTGCGCGCGCGCGCGCGG 4747
 QY 103 TTCTAGGCGCGCACAGTGCAGAGATCTCTCGCGGGCTCGGTACCAAGGCGCGGATG 162
 Db 4746 ATCTCTCGCGCGC - TGGGGCGGTGTTCGCGCGCGCGCTTCCCGCGCGCGCGCGAGAC 4688
 QY 163 TACTTTCACCTTCTCCAAAGAGAGCTGGCGCGCGGTGTGTGCGCGCGAGCAGACCTTG 222
 Db 4687 CCAGCGCGCGCGCGCACCGCGCGAGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCG 4628
 QY 223 CAGGTGGCGGTGCGGATTCGGCTCAAGGCGCGAGAACTGGTAGACCTCAACATGCTG 282
 Db 4627 CCGCTGGCGGTGAGTGGCTTGGCGCGCTTCCCGCGGATCTTGGAGCGGTGGCGAGGCG 4568
 QY 283 GTGCGCGCGCGATGTCGACGATCGATCTTGGCGCGCGCGCGCGCGCGCGCGCGCG 342
 Db 4567 TTGAGCGCGAGCTTGGCGCGCTTCCCGGGGTGCGCGCGCGCGCGCGCGCGCGCG 4508
 QY 343 CAGGGGCGGTGGACTTCTCCGACCGCAACCCGTTCGCGCGAGTGGGGCGGACATCTGCG 402
 Db 4507 CGGCGCGAGGACCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCG 4448
 QY 403 CAGCTCTCGCGAGCAGAGAACCGGGGAGGTGCTTCCGCACTGTAACCGGAAAG 462
 Db 4447 CTGCGCGGTGCTGCGCGAGCTGCGGTTCGTGCGCGAGCGCGGTGCTCATGCGCG 4388
 QY 463 ACCGGGACTTTCATGTCGGGTGCTTACCGCGGCTCCAGCGGTCTCCGGGTCACTCC 522
 Db 4387 CCGGGGACTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4328
 QY 523 GACCGCGAGACTCGGCGCACCGGATCTCGGTGATGTGGAACCACTGCTGCGCGAGCATC 582
 Db 4327 AGCTTGTGCGCGCGCGCGCTGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCG 4268
 QY 583 GTGCGCGGTCTCATGTGACTGATCGAATCGGCGGAGAGCGGATCGGAGGTGCGG 642
 Db 4267 GCG 4208
 QY 643 GCGGGGCG 693
 Db 4207 GCGGGGCG 4157

RESULT 13
 AAS54248
 ID AAS54248 standard; DNA; 1041 BP.
 AC AAS54248;
 XX
 DT 13-FEB-2002 (first entry)
 DE Pseudomonas aeruginosa DNA for cellular proliferation protein #379.
 KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.

XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 DR P-PSDB; AAU36389.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Claim 27; SEQ ID NO 7885; 511bp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1041 BP; 157 A; 347 C; 385 G; 152 T; 0 U; 0 Other;

Query Match 9.6%; Score 67.2; DB 4; Length 1041;
 Best Local Similarity 46.9%; Pred. No. 0.0097;
 Matches 210; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
 QY 212 AGCAGACCTTGACATGCGCGGTGCGCGGTCCAAAGCGCGAGGAGTGTTCGCGAGACC 271
 Db 575 AGCGCGCATCGAGCTGCGCGCGCGCTATGGCGAGGTGATGCTGGAGCGCTTCGTCG 634
 QY 272 TCACATGCTGTGTCGCCACGCGATGCTGACGATCCGATCTGCGGGCGGCGACGCGGC 331
 Db 635 CCGGCGCGAGGTACCGCTCGGGGTGCTCGACGACGCGCTGCGGTGGCGGAGATTC 694
 QY 332 TCGCACTGACCGAGCGCGGTGGACTTCTCCGACGCCAACCCGTTCCGCGAGTGGGGG 391
 Db 695 TCTCGCGCGCGAGAGGTCTTCGACTAGAGCAAGTACAGCGCGCGCGGTGCGCG 754
 QY 392 ACATCTGCGCGCGAGCTCTTGGCGGAGGACAGAGAACGGGGGAGGTGTTCGCGACGTGA 451
 Db 755 AGGTGTTTCCCGCGCGACCTGCGCGCGCGATCGCGCGAGGCCCGAGCGCTGCGCTGA 814
 QY 452 ACCCGAAAAGACCGCGGACTTCATGCTGCGGTGCTTACCGGGGTCCAGGCGGTCTCC 511
 Db 815 AGGTACACCGGGCGCTGAAGCTGAGCGGTACAGCGCGACCGACTTCCCGCTCGACGAAC 874
 QY 512 GGGTCACCTCCGACCGCGAGGACCTCGGCCACCGGATCTCGGTGATGTGGAACACGTCG 571
 Db 875 AGGGCGGCTCTGTGTGCTTGGAGGTCAATACCTCGCGGGGATGACCGCCACGAGCTGC 934
 QY 572 TGCCCGAGCATGTCGCGCGGTCCATGCTGACCTGGATCGAAACCGCGGAGGAGCGGATCG 631
 Db 935 TGCGCGAGCGCGCGCGCGCGCGGTGCGGTTCGCGGAACTCTCGAGCGGATCTGCGC 994
 QY 632 GGAAGTTCG 659
 Db 995 GGCCTCGCATCGAGCGCTGCAAGGCGCG 1022

RESULT 14
 ACA42557
 ID ACA42557 standard; DNA; 1041 BP.

XX AC A42557;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #24214.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Pseudomonas aeruginosa.
XX PN W0200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX DR P-PSDB; ABU38687.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 30427; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1041 BP; 157 A; 347 C; 385 G; 152 T; 0 U; 0 Other;

Query Match

9.6%; Score 67.2; DB 8; Length 1041;

Best Local Similarity 46.9%; Pred. No. 0.0097;
Matches 210; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
QY 212 AGCAGACCTGACAGTGGCGGTGCGGAATCGGGCTCCAGGCGCAGGAACCTGGTAGACC 271
DB 575 AGCGGCCCATCGAGCTGGCGCGGCTATGGGCACAGGTGATGCTGGAGCGCTTGTGTCG 634
QY 272 TCACCATGCTGGTTCGCCACCGCATCTGCACGATCCGATCTCTGCGGGCGGCACGCGGC 331
DB 635 CCGGCCGCGAGGTTCACCGTCGGGTCTCGACGACGAGCGCTGCGGTGGCGGAGATTC 694
QY 332 TCGCACTGACAGGCGGTGGACTTCTCCGACCCCAACCCGTTCCGGAGTGGGGCG 391
DB 695 TCTCGCGGCCAGGAGGTCTTCGACTACGAGCACAAATACCAAGCGCGCGGTGCGCG 754
QY 392 ACATCTGCGCCCGAGCTCTTGGCGGAGGCACAGAAACGGGGGAGGTGCTTCCCGACGTGA 451
DB 755 AGGTGTTTCCCGCGGACCTGCGCGCGCGATGCGCGCGAGGCCCGCTGCGCGCTGA 814
QY 452 ACCCGAAAAGACCGGCGACTTCATCGTGGGTGCTTACCGGGCTCCAGGCGGTCTCC 511
DB 815 AGGTACACCGGCGCTGAAAGCTGAGCGGCTACAGCGCACCGACTTCCGCGCTCGACGAAC 874
QY 512 GGTACACTCCGACCGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGGAACACGTCG 571
DB 875 AGGGCGCGCTCTGTGCTGAGGTCAATACCTCGCGGCGATGACCGCCACGAGCTGC 934
QY 572 TGCCCAAGCATCTGCGCGGCTCCATGCTGACCTGGAATCGAAACCGCGAGGAGCGGATCG 631
DB 935 TCGCGGAGCGCGCGCGCGCGGATCGGTTCGCCGAACTCTCGAGCGGATCTGCC 994
QY 632 GGAAGTTCGCGCGCGCGCGCGGCGCG 659
DB 995 GGTGCGCATCGAGCGCTCAAGGGCGC 1022
RESULT 15
ABD05332
ID ABD05332 standard; DNA; 1107 BP.
XX AC ABD05332;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polynucleotide #3936.
XX KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX KW antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nelling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
XX DR P-PSDB; AB071761.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 3936; 455pp; English.

us-10-049-710a-1.rng

The invention relates to *Pseudomonas aeruginosa* polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective anti-bacterial targets, as targets for antibacterial drugs, including anti-*P. aeruginosa* drugs, as templates for recombinant production of *P. aeruginosa*-derived peptides or polypeptides, as target components for diagnosis and/or treatment of *P. aeruginosa*-caused infection, and in detection of *P. aeruginosa* sequences or other sequences of *Pseudomonas* species using biochip technology. Sequences ABD01397-ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 1107 BP; 168 A; 366 C; 401 G; 172 T; 0 U; 0 Other; XX SO

Query Match 9.4%; Score 66; DB 11; Length 1107;

Best Local Similarity 47.6%; Pred. No. 0.016;

Best local similarity: 100%; Identical: 100%;
Matches 195; Conservative 0; Mismatches 215; Indels 0; Gaps 0;

QY 212 AGCAGACCCCTGCACGTGGCGGTCCCGGAATCCGGCTCCAAGGCGCAGGAACCTGGTAGACC 271

641 AGCCGGCCATCGAGCTGGCCGGCGCTATGGCGACGAGGTGATGCTGGAGCGCTTCGTCG 700

QY 272 TCACCATGCTGGTCGCCACGGCATGCTGCACGATCCGATCCTGCGGGCGGCACGCGGC 331

db
701 CCGGCCGAGGTACCGTGGGGTCTCGACGACGAGCGCTCCGGTGGCGGAGATTC 760

332 TCGCACTGGACCAGGGGGCGGTGGACTTCTCCGACGCCAACCCGTTCCGGCGAGTGCGGCG 391

Db
761 TCCTCGGCGCCAGGAGGCTTCGACTACGAGCACAAAGTACCAGCGCGCGGTGCGCG 820

392 ACATCTGCGCCAGCTCCTGGCGGAGGCACAGGAACGGGGGAGGTGCTTCCGCACGTGA 451

821 AGGTGTTTCCCGCCGACCTGCCGCCGCGATCGCCGCCGAGGCCAGCGTCTGGCGCTGA 880

452 ACCCGAAAAAGACCGGGACTTCATCGTCGGCTGCTTACCGGGCTCCAGGCGGTCCTCCC 511

881 AGGTACACCGGGCGCTGAAGCTGAGCGGTACAGCCGACCGACTTCCGCCCTCGACGAAC 940

512 GGGTCACCTCCGACCGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGGAACCACTGC 571

941 AGGGCCGGCTCTGGTGCCTGGAGGTCAATACCTGCCGGCATGACCGCCACCAAGCTGC 1000

572 TGCCCAGCATCGTGCCGGGTCCATGCTGACCTGGATCGAAACCGCGAG 621

db 1001 TGCCGAGCCGCGCGCGCGGATCGGTTTCGCCAACTCTGCAG 1050

Search completed: March 6, 2005, 20:06:59

Job time : 594 secs

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:54:40 ; Search time 227 Seconds
(without alignments)
5038.579 Million cell updates/sec

Title: US-10-049-710A-1
Perfect score: 699
Sequence: 1 atggcagtcgacagcaagc.....aggccgctccagcagtag 699

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.2	11.2	4257	2	US-08-690-473-1
2	78.2	11.2	4257	3	US-09-259-821A-1
3	78.2	11.2	4257	3	US-08-843-659-1
4	78.2	11.2	4257	4	US-09-825-288A-1
5	67.8	9.7	12001	1	US-08-458-568A-11
6	66	9.4	1107	4	US-09-252-991A-3936
7	66	9.4	1134	4	US-09-252-991A-3908
8	66	9.4	1326	4	US-09-252-991A-4002
9	65	9.3	77536	4	US-09-410-551B-1
10	65	9.3	77536	4	US-09-940-316B-1
11	62.8	9.0	4403765	3	US-09-103-840A-2
12	62.8	9.0	4411529	3	US-09-103-840A-1
13	62	8.9	630	4	US-09-302-540-4889
14	62	8.9	28493	4	US-09-302-540-1241
15	60.2	8.6	23673	3	US-09-773-816-1
16	59.6	8.5	2721	6	5215881-2
17	59.6	8.5	2721	6	5215881-2
18	58.4	8.4	1536	4	US-09-302-540-8768
19	58.4	8.4	9185	4	US-09-302-540-931
20	58.2	8.3	3957	4	US-10-237-551-193
21	58.2	8.3	154746	4	US-09-827-688-8
22	58.2	8.3	154746	4	US-09-827-688-8
23	58	8.3	8438	1	US-07-945-283-1
24	58	8.3	15872	3	US-09-105-537-1
25	58	8.3	15872	4	US-09-091-609-1
26	58	8.3	15872	4	US-09-091-609-3
27	57.2	8.2	1626	3	US-09-158-767-14

28	57.2	8.2	1626	4	US-09-713-794-14	Sequence 14, Appl
29	57.2	8.2	2181	3	US-09-158-767-10	Sequence 10, Appl
30	57.2	8.2	2181	4	US-09-713-794-10	Sequence 10, Appl
31	57	8.2	34662	4	US-09-902-540-1261	Sequence 1261, Ap
32	56.8	8.1	717	4	US-09-252-991A-11242	Sequence 11242, A
33	56.8	8.1	1632	4	US-09-252-991A-11015	Sequence 11015, A
34	56	8.0	1197	4	US-10-151-832-2	Sequence 2, Appl
35	56	8.0	1227	4	US-09-252-991A-16560	Sequence 16560, A
36	56	8.0	1308	4	US-10-151-832-1	Sequence 1, Appl
37	56	8.0	1368	4	US-09-252-991A-16040	Sequence 16040, A
38	56	8.0	2987	3	US-09-105-537-40	Sequence 40, Appl
39	56	8.0	5787	3	US-09-320-878-21	Sequence 21, Appl
40	56	8.0	5970	4	US-09-141-908-11	Sequence 11, Appl
41	56	8.0	5970	4	US-09-657-440-21	Sequence 21, Appl
42	55.6	8.0	378	4	US-09-902-540-2074	Sequence 2074, Ap
43	55.6	8.0	2028	4	US-09-252-991A-13079	Sequence 13079, A
44	55.6	8.0	3525	4	US-09-252-991A-12708	Sequence 12708, A
45	55.6	8.0	3906	4	US-09-252-991A-13251	Sequence 13251, A

ALIGNMENTS

RESULT 1
US-08-690-473-1
; Sequence 1, Application US/08690473
; Patent No. 5876923
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN
; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/690,473
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD:239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-690-473-1

Query Match 11.2%; Score 78.2; DB 2; Length 4257;
Best Local Similarity 45.0%; Pred. No. 1.3e-07;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
OY 43 GCGTCGCGCAGCGGCGAGTCGTGCGGGGAGCGCCTCGGTCTTCGACAGTACGGG 102
Db 1867 GCGTCGCGCAGCGGCGAGTCGTGCGGGGAGCGCCTCGGTCTTCGACAGTACGGG 1926
OY 103 TTCGAGCGCCGACAGTGGCAGGATCCTTCGCGGGCCTCGGTTCACCAAGGGCGGATG 162

Db 1927 ATCTCGCGCCCTGGGGCGGCTGTCTCGCGCGCCCGCTCCCGCGGGGGCGACGAC 1986
QY 163 TACTTCACTTCCGTTCCAAAGGAGAGCTGGCGCGCGCTGCTGGCGGAGCAGACCTTG 222
Db 1987 CCGGAGCGCGCGCGGCAAGCGGAGAGAGAGCGCGGCGCGCGCGCGCGCGCGCGCG 2046
QY 223 CAGGTGGCGGTGCGGAAATCCGGCTCCAAAGGCGGAGGAACTGTAGACCTCAACATGCTG 282
Db 2047 CGGTGGCGGTGAGTGGCTGGCGCGCTGCGCGCGGATCTTGGAGCGCTGGCGGAGGC 2106
QY 283 GTGCGCCACGGCATGCTGCAAGATCCGATCTGCGCGGCGGCGACGCGGCTGCGACTGAG 342
Db 2107 TTCGACGGGACCTGGCGCGCTGCGCGCGCTGCGCGCGGCTGCGCGCGGCGCGCG 2166
QY 343 CAGGGGCGGTGAGTCTTCCGACGCAACCGGCTTCCGCGAGTGGGGCGACATCTGCGCC 402
Db 2167 CGCGCGGAGGACCGCGGGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTTGGCGGAGGACAGGAAACGGGGGAGGTGCTTCCGCGAGTGAACCGGAAAG 462
Db 2227 CTGCGCGCTGGCTGCGGAGCTGCGGCTTCTGCGCGAGCGGCTGTGCTCATGCGCTG 2286
QY 463 ACGGCGGACTTATCTGCGGTCTTCAACGGGCTCCAGGCGTCTCCGGGTCACTCC 522
Db 2287 CGCGGGGAGCTGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2346
QY 523 GACCGCGAGGACCTCGGCGCACCGGATCTCGGTGATGTGAACACGCTGTGCGCGCATC 582
Db 2347 AGCTGTGCGCGGGCGCTTGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCG 2406
QY 583 GTGCGCGCTCATGCTGACCTGGAATCGAAACCGGCGAGGCGGATCGGGAGGTGCGG 642
Db 2407 GCG 2466
QY 643 GCG 693
Db 2467 GCG 2517

RESULT 2

US-09-259-821A-1

; Sequence 1, Application US/09259821A

; Patent No. 6210926

; GENERAL INFORMATION:

; APPLICANT: LEOPARDI, ROSARIO

; APPLICANT: ROIZMAN, BERNARD

; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS

; FILE REFERENCE: ARCD:317

; CURRENT APPLICATION NUMBER: US/09/259,821A

; CURRENT FILING DATE: 1999-03-01

; PRIOR FILING DATE: 08/690,473

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 4257

; TYPE: DNA

; ORGANISM: HERPES VIRUS, TYPE 1

US-09-259-821A-1

Query Match

Best Local Similarity 11.2%; Score 78.2; DB 3; Length 4257;

Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY 43 GCGTCCGCGCGCGCGCGGATCGTGGGCGAGCGCGCTCGTTCGACGAGTACGGG 102
Db 1867 GCGTCCG 1926
QY 103 TTGAGGCG 162
Db 1927 ATCTCG 1986

QY 163 TACTTCACTTCCGTTCCAAAGGAGAGCTGGCGCGCGCTGCTGGCGGAGCAGACCTTG 222
Db 1987 CCGGAGCGCGCGCGGCAAGCGGAGAGAGAGCGCGGCGCGCGCGCGCGCGCGCG 2046
QY 223 CAGGTGGCGGTGCGGAAATCCGGCTCCAAAGGCGGAGGAACTGTAGACCTCAACATGCTG 282
Db 2047 CGGTGGCGGTGAGTGGCTGGCGCGCTGCGCGCGGATCTTGGAGCGCTGGCGGAGGC 2106
QY 283 GTGCGCCACGGCATGCTGCAAGATCCGATCTGCGCGGCGGCGACGCGGCTGCGACTGAG 342
Db 2107 TTCGACGGGACCTGGCGCGCTGCGCGCGCTGCGCGCGGCTGCGCGCGGCGCGCG 2166
QY 343 CAGGGGCGGTGAGTCTTCCGACGCAACCGGCTTCCGCGAGTGGGGCGACATCTGCGCC 402
Db 2167 CGCGCGGAGGACCGCGGGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTTGGCGGAGGACAGGAAACGGGGGAGGTGCTTCCGCGAGTGAACCGGAAAG 462
Db 2227 CTGCGCGCTGGCTGCGGAGCTGCGGCTTCTGCGCGAGCGGCTGTGCTCATGCGCTG 2286
QY 463 ACGGCGGACTTATCTGCGGTCTTCAACGGGCTCCAGGCGTCTCCGGGTCACTCC 522
Db 2287 CGCGGGGAGCTGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2346
QY 523 GACCGCGAGGACCTCGGCGCACCGGATCTCGGTGATGTGAACACGCTGTGCGCGCATC 582
Db 2347 AGCTGTGCGCGGGCGCTTGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 2406
QY 583 GTGCGCGCTCATGCTGACCTGGAATCGAAACCGGCGAGGCGGATCGGGAGGTGCGG 642
Db 2407 GCG 2466
QY 643 GCG 693
Db 2467 GCG 2517

RESULT 3

US-08-843-659-1

; Sequence 1, Application US/08843659

; Patent No. 6218103

; GENERAL INFORMATION:

; APPLICANT: Leopardi, Rosario

; APPLICANT: Roizman, Bernard

; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS

; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: United States

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/843,659

; FILING DATE: Concurrently Herewith

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Highlander, Steven L.

; REGISTRATION NUMBER: 37,642

; REFERENCE/DOCKET NUMBER: ARSB:519

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (512) 474-7577

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4257 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-843-659-1

Query Match      11.2%; Score 78.2; DB 3; Length 4257;
Best Local Similarity 45.0%; Pred. No. 1.3e-07;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY 43 GCGCTCCGACGCGGAGCGATCGTGGGAGCGCGCTCGTCTTCACGAGTACGGG 102
DB 1867 GCGTTCGCGCGCGCCACCGGCGAGCGCGGTTCGCCCGGTTACGGCGCGCGGG 1926
QY 103 TTCGAGGCGCGCACATGCGAGATCTCTCGCGGCGCTCGGTTCACCAAGGCGCGATG 162
DB 1927 ATCTCGCGCGCGCTGGGCGGTGTCTCGCGCGCGCTCTCCCGCGGCGGCGACGAC 1986
QY 163 TACTTCCACTTCGCTTCCAAAGAGAGCTGGCGCGGCGGTGTGGCGGAGACCCCTG 222
DB 1987 CCGAGCGCGCGCGCCACCGCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 2046
QY 223 CAGCTGGCGGTTCGCGAATCCGCGTCCAAAGCGCGAGAACTGTAGACCTCACCATGCTG 282
DB 2047 CGCGTGGCGGTTCGAGTGCCTGCGCGCTGCGCGGATCTTGAGGCGCTGGCGGAGGC 2106
QY 283 GTGCGCCACGGATGTCGACGATCCGATCTCGCGGCGGCGACGCGGCTGCGACTGGAC 342
DB 2107 TTCGACGGCGACCTGGCGCGCTTCCCGGCGCTGGCGCGGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGCGGTGAGATCTTCGACGCGCAACCGCTTCGCGGAGTGGGCGGACATCTCGGC 402
DB 2167 CGCGCGGAGGAGCCCGCGGCGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTCTGGCGAGGACAGAGAACCGGGGAGGTCTTCGACGCTGAACCGGAAAG 462
DB 2227 CTGCGCGGTGCTGCGCGAGCTGCGTTCGTCGCGAGCGCTGTGTCTCATGCGCTG 2286
QY 463 ACCGGGACTTCATGTGCGGTCTGTTACCGGGCTCCAGCGGCTTCGCGGTTCACTCC 522
DB 2287 CGCGGGGACCTGCGGTGCGCGGCGGACGAGCGCGCGCTGGCGCGCGCTGGCGCG 2346
QY 523 GACCGCGAGGACTCGGCGCACCGGATCTCGGTGATGTGAACACGCTGTGCCAGCATC 582
DB 2347 AGCTTGGTCCGCGGCGCTTCGCGCGCGCTTCGCGCGGAGCCCGCGCTTCGAGCTCC 2406
QY 583 GTGCGCGCTTCATGCTGACCTGATTCGAAACCGCGGAGGCGGATCGGGAAGTCCGG 642
DB 2407 GCGCGCGCGCGCGCGCGGACCTGCTGTTTGAACAACAGAGCTGCGCGCGCTGCTGGCG 2466
QY 643 GCGCGCGCGCGCGCGCGCGGAGCTGCGGAGGCTTCGAGGCGCGCTTCGAC 693
DB 2467 GCGCGCGCGCGCGCGCGCGGAGCGCGCGCGCTGGCGCGCGCGCGCGCGCG 2517
```

```
RESULT 4
US-09-825-288A-1
; Sequence 1, Application US/09825288A
; Patent No. 6723511
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317USC1
; CURRENT APPLICATION NUMBER: US/09/825,288A
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 09/259,821
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
```

```
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-1

Query Match      11.2%; Score 78.2; DB 4; Length 4257;
Best Local Similarity 45.0%; Pred. No. 1.3e-07;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY 43 GCGCTCCGACGCGGAGCGATCGTGGGAGCGCGCTCGTCTTCACGAGTACGGG 102
DB 1867 GCGTTCGCGCGCGCCACCGGCGAGCGCGGTTCGCCCGGTTACGGCGCGCGGG 1926
QY 103 TTCGAGGCGCGCACATGCGAGATCTCTCGCGGCGCTCGGTTCACCAAGGCGCGATG 162
DB 1927 ATCTCGCGCGCGCTGGGCGGTGTCTCGCGCGCGCTCTCCCGCGGCGGCGACGAC 1986
QY 163 TACTTCCACTTCGCTTCCAAAGAGAGCTGGCGCGGCGGTGTGGCGGAGACCCCTG 222
DB 1987 CCGAGCGCGCGCGCCACCGCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 2046
QY 223 CAGCTGGCGGTTCGCGAATCCGCGTCCAAAGCGCGAGAACTGTAGACCTCACCATGCTG 282
DB 2047 CGCGTGGCGGTTCGAGTGCCTGCGCGCTGCGCGGATCTTGAGGCGCTGGCGGAGGC 2106
QY 283 GTGCGCCACGGATGTCGACGATCCGATCTCGCGGCGGCGACGCGGCTGCGACTGGAC 342
DB 2107 TTCGACGGCGACCTGGCGCGCTTCCCGGCGCTGGCGCGGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGCGGTGAGATCTTCGACGCGCAACCGCTTCGCGGAGTGGGCGGACATCTCGGC 402
DB 2167 CGCGCGGAGGAGCCCGCGGCGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTCTGGCGAGGACAGAGAACCGGGGAGGTCTTCGACGCTGAACCGGAAAG 462
DB 2227 CTGCGCGGTGCTGCGCGAGCTGCGTTCGTCGCGAGCGCTGTGTCTCATGCGCTG 2286
QY 463 ACCGGGACTTCATGTGCGGTCTGTTACCGGGCTCCAGCGGCTTCGCGGTTCACTCC 522
DB 2287 CGCGGGGACCTGCGGTGCGCGGCGGACGAGCGCGCGCTGGCGCGCGCTGGCGCG 2346
QY 523 GACCGCGAGGACTCGGCGCACCGGATCTCGGTGATGTGAACACGCTGTGCCAGCATC 582
DB 2347 AGCTTGGTCCGCGGCGCTTCGCGCGCGCTTCGCGCGGAGCCCGCGCTTCGAGCTCC 2406
QY 583 GTGCGCGCTTCATGCTGACCTGATTCGAAACCGCGGAGGCGGATCGGGAAGTCCGG 642
DB 2407 GCGCGCGCGCGCGCGCGGACCTGCTGTTTGAACAACAGAGCTGCGCGCGCTGCTGGCG 2466
QY 643 GCGCGCGCGCGCGCGCGGAGCTGCGGAGGCTTCGAGGCGCGCTTCGAC 693
DB 2467 GCGCGCGCGCGCGCGCGGAGCGCGCGCGCTGGCGCGCGCGCGCGCGCGCG 2517
```

```
RESULT 5
US-08-458-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; APPLICANT: Yeh, Lily
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
; TITLE OF INVENTION: Infections
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,568A
FILING DATE: 02-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/065,146
FILING DATE: 05-MAY-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn R.
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: DFCI-0029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 12001 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Herpes simplex virus
STRAIN: Herpes Simplex Virus Type 1
US-08-458-568A-11

Query Match 9.7%; Score 67.8; DB 1; Length 12001;
Best Local Similarity 45.0%; Pred. No. 2.4e-05;
Matches 293; Conservative 0; Mismatches 357; Indels 1; Gaps 1;
QY 43 GCCGTCGACGCGGAGCGATCTGTGGGAGCGCGCTCTTTCGACGAGTACGGG 102
DB 4806 GCCGTGCGCGCGCGACCGCGAGCGCGGTGTCGCGGTCGCGCGCGCGGGG 4747
QY 103 TTGAGGCGCGCACAGTGCAGAGATCTCTCGGGGCTTCGTCACCAAGGCGCGATG 162
DB 4746 ATCTCGCGCGCC-TGGGGCGGTGTCTCCGCGCGCGCTCTCCCGCGGGGCGACGAC 4688
QY 163 TACTTCCACTTCGCTTCCAAAGAGCTGGCGCGCGGTGTGGCGGAGCAGACCTCTG 222
DB 4687 CCGGACGCGCGCGCACCGCACCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 4628
QY 223 CACGTGGCGGTGCGGAAATCCGGCTCCAAAGGCGCAGGAATGCTAGACTCACTCATGTG 282
DB 4627 CGCGTGGCGGTGCGAGTGCCTGCGCGCTGCGCGGATCTGGAGCGCTGGCGGAGGCG 4568
QY 283 GTGCGCCAGCGCATGCTGACAGTCCGATCCGCGGCGGCGACGCGGCTGCGCATGGAC 342
DB 4567 TTGAGCGGCGACTGCGCGCGCTGCGCGGCGGTGCGCGGCGCGCGCGCGCGCGCG 4508
QY 343 CAGGGGCGGTGGAGTCTTCCGACGCGCAACCGCTTTCGGGAGTGGGGCGACATCTCGCGC 402
DB 4507 CGGCGGAGGAGCCCGCGGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 4448
QY 403 CAGTCTCTGGCGAGGCA CAGGAA CCGGGGAGGTGCTTTCGCGCATGTAACCCGAAAG 462
DB 4447 CTGCGCGGTGCGTGGCGAGTGTGCTGTCGCGCGAGCGCTGTGCTCATGCGCCTG 4388
QY 463 ACCGGGACTTCATGTCGGGTGCTTACCGGGCTCCAGGCGGTCTCCGGGTCACTCC 522
DB 4387 CCGGGGACTCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4328
QY 523 GACCGCGAGACTCTCGGCGACCGGATCTCGGTGATGTGGAACCACTGCTGCCCGCATC 582
DB 4327 AGCTGTCTCGGGGCGCTTGGCGCGCGCGCTGCGCGGAGCCCGCGCTTCCGAGTCC 4268
QY 583 GTGCGGCGCTTCATGCTGACTGATCGAAGAACCGGCGAGGAGCGGATCGGGAGGTGCGG 642

DB 4267 GCGGCGCGCGCGCGCGCGAGCTGCTTTTGACACAGAGAGCTGCGCGCCCTGCTGGCG 4208
QY 643 GCGGCGCGCGAGGCGCGGAGGCTCGGAGGCTCGGAGGCGCTCCGAC 693
DB 4207 GCGGCGCGCGAGCGCACCGGACCGCGCGCTGGCGCGCGCGCGCGCG 4157
RESULT 6
US-09-252-991A-3936
Sequence 3936, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3936
LENGTH: 1107
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3936

Query Match 9.4%; Score 66; DB 4; Length 1107;
Best Local Similarity 47.6%; Pred. No. 4.1e-05;
Matches 195; Conservative 0; Mismatches 215; Indels 0; Gaps 0;
QY 212 AGCAGACCTGACGCTGCGGTGCGGATCCGGTCCAGCGCAGGAACTGGTAGACC 271
DB 641 AGCGCGCATCGAGCTGCGCGCGCTATGCGCAGAGGTGATGTGAGCGCTTCGTG 700
QY 272 TCACCATGCTGTGCGCCACGCGCATCTGACGATCCGATCTGCGGGCGGCGACGCGC 331
DB 701 CGGCGCGGAGGTGACCGTTCGGGTGCTCGACGACGAGCGCTGCGGTGGCGAGATTC 760
QY 332 TCGCACTGAGACAGGGGCGGTGGAATCTTCGACGCAACCCGTTTCGCGAGTGGGGCG 391
DB 761 TCTCGCGCGCGCAGGAGGTCTTCGACTACGAGCACAAGTACCGCGCGCGGTGCGCG 820
QY 392 ACATCTGCGCGCGAGCTCTGCGGAGGACAGGACGCGGGGAGGTGCTTCGCGACGTGA 451
DB 821 AGGTGTTTCCCGCGCGACTGCGCGCGCGATCGCGCGGAGCCAGCGGTCTGGCGCTGA 880
QY 452 ACCCGAAAAAGACCGCGACTTTCATCTGCTGCGGTGCTTACCGGGCTCCAGGGGCTCC 511
DB 881 AGGTACACGCGCGCTGAGCTGAGCGGTACAGCGGACCGACTTCGCGCTCGAGGAC 940
QY 512 GGGTCACCTCGACCGCGCAGGACTTCGGCCACCGGATCTCGGTGATGTGGAACCAAGTGC 571
DB 941 AGGGCGCGCTCTGGTGCCTGAGGTCAATACCTCGGGGATGACCGCGCACGAGCTGC 1000
QY 572 TGCCAGCATGCTGCGCGGTTCATCTGACTGATCGAATCGGCGAG 621
DB 1001 TGCCGAGCGCGCGCGCGCGCGGATCGGTTTCGCGAACTCTGCGAG 1050

RESULT 7
US-09-252-991A-3908
Sequence 3908, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3908
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3908

Query Match          9.4%; Score 66; DB 4; Length 1134;
Best Local Similarity 47.6%; Pred. No. 4.1e-05;
Matches 195; Conservative 0; Mismatches 215; Indels 0; Gaps 0;

QY 212 AGCAGACCTTGCACGTGGCGGTGCCGAATCCGGCTCCAAAGCGCACGGAACCTGGTAGACC 271
DB 212 AGCAGACCTTGCACGTGGCGGTGCCGAATCCGGCTCCAAAGCGCACGGAACCTGGTAGACC 271
QY 272 TCACCATGCTGTCGCCCAAGGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 331
DB 272 TCACCATGCTGTCGCCCAAGGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 331
QY 585 CGCGCGCGAGGTCAACCGTGGGGTGTCTGACGACCAAGGCTGCGGCTGGCGGTGGCGGATTC 644
DB 585 CGCGCGCGAGGTCAACCGTGGGGTGTCTGACGACCAAGGCTGCGGCTGGCGGTGGCGGATTC 644
QY 332 TGGCACTGACGAGGCGGCTGGGCTCTCCAGCCCAACCGCTTCCGGCGAGTGGGGCG 391
DB 332 TGGCACTGACGAGGCGGCTGGGCTCTCCAGCCCAACCGCTTCCGGCGAGTGGGGCG 391
QY 645 TCTCGCGCGCAGGAGGCTTCTGACTAGACCAAGTACCAGCGCGCGCGGTGGCGG 704
DB 645 TCTCGCGCGCAGGAGGCTTCTGACTAGACCAAGTACCAGCGCGCGCGGTGGCGG 704
QY 392 ACATCTGCCGCCAGCTCTTGGCGGAGGACAGAAACGGGGGAGGTGCTTCCGCGAGTGA 451
DB 392 ACATCTGCCGCCAGCTCTTGGCGGAGGACAGAAACGGGGGAGGTGCTTCCGCGAGTGA 451
QY 705 AGGTGTTTCCCGCGACCTGCGCGCGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGTGA 764
DB 705 AGGTGTTTCCCGCGACCTGCGCGCGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGTGA 764
QY 452 ACCCGAAAAGACCGCGGACTTCATCGTGGGTGCTTCAACCGGGCTTCAAGGGGTCTCC 511
DB 452 ACCCGAAAAGACCGCGGACTTCATCGTGGGTGCTTCAACCGGGCTTCAAGGGGTCTCC 511
QY 765 AGGTACACCGGGCGGTGAAGCTGAGCGGTACAGCGCGCACCGACTTCCGCGCTCGAGAAC 824
DB 765 AGGTACACCGGGCGGTGAAGCTGAGCGGTACAGCGCGCACCGACTTCCGCGCTCGAGAAC 824
QY 512 GGTACCTCCGACCGCGAGGCTCGGCCACCGATCTCGGTGATGTGGAACCGAGTGC 571
DB 512 GGTACCTCCGACCGCGAGGCTCGGCCACCGATCTCGGTGATGTGGAACCGAGTGC 571
QY 825 AGGGCGGCTTGGTGGCTTGGAGGTCAATACCTGCGGGCGATGACCGCGCGAGCTGC 884
DB 825 AGGGCGGCTTGGTGGCTTGGAGGTCAATACCTGCGGGCGATGACCGCGCGAGCTGC 884
QY 572 TGCCAGCATCTGTCGGCGGTTCATCTGCTGATCGATCGATCGATCGATCGATCGATCGAT 621
DB 572 TGCCAGCATCTGTCGGCGGTTCATCTGCTGATCGATCGATCGATCGATCGATCGATCGAT 621
QY 885 TGCCGAGCGCGCGCGCGCGCGGTTCGCGGATCGGTTCGCGGACTCTGCGAG 934
DB 885 TGCCGAGCGCGCGCGCGCGCGGTTCGCGGATCGGTTCGCGGACTCTGCGAG 934

RESULT 8
US-09-252-991A-4002/c
; Sequence 4002, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4002
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4002

Query Match          9.4%; Score 66; DB 4; Length 1326;
Best Local Similarity 47.6%; Pred. No. 4.2e-05;
Matches 195; Conservative 0; Mismatches 215; Indels 0; Gaps 0;

QY 212 AGCAGACCTTGCACGTGGCGGTGCCGAATCCGGCTCCAAAGCGCACGGAACCTGGTAGACC 271
DB 212 AGCAGACCTTGCACGTGGCGGTGCCGAATCCGGCTCCAAAGCGCACGGAACCTGGTAGACC 271
QY 506 AGCCGCCATCGAGCTGGCGGCGGTATGGGACAGAGTGATCTGGAGCGGCTTCTGTCG 447
DB 506 AGCCGCCATCGAGCTGGCGGCGGTATGGGACAGAGTGATCTGGAGCGGCTTCTGTCG 447
```

```
QY 272 TCACCATGCTGTCGCCCAAGGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 331
DB 272 TCACCATGCTGTCGCCCAAGGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 331
QY 446 CGCGCGCGAGGTCAACCGTGGGGTGTCTGACGACCAAGTACCAGCGCGCGGTGGCGGATTC 387
DB 446 CGCGCGCGAGGTCAACCGTGGGGTGTCTGACGACCAAGTACCAGCGCGCGGTGGCGGATTC 387
QY 332 TGGCACTGACGAGGCGGCTGGGCTCTCCAGCCCAACCGCTTCCGGCGAGTGGGGCG 391
DB 332 TGGCACTGACGAGGCGGCTGGGCTCTCCAGCCCAACCGCTTCCGGCGAGTGGGGCG 391
QY 386 TCTCGCGCGCGAGGAGGTCTTTCGACTAGACCAAGTACCAGCGCGCGGTGGCGG 327
DB 386 TCTCGCGCGCGAGGAGGTCTTTCGACTAGACCAAGTACCAGCGCGCGGTGGCGG 327
QY 392 ACATCTGCCGCCAGCTCTTGGCGGAGGACAGAAACGGGGGAGGTGCTTCCGCGAGTGA 451
DB 392 ACATCTGCCGCCAGCTCTTGGCGGAGGACAGAAACGGGGGAGGTGCTTCCGCGAGTGA 451
QY 326 AGGTGTTTCCCGCGACCTGCGCGCGCGATCGCGCGCGCGCGCGCGCGCGCGCGTGA 267
DB 326 AGGTGTTTCCCGCGACCTGCGCGCGCGATCGCGCGCGCGCGCGCGCGCGCGCGTGA 267
QY 452 ACCCGAAAAGACCGCGGACTTCATCGTGGGTGCTTCAACCGGGCTTCAAGGGGTCTCC 511
DB 452 ACCCGAAAAGACCGCGGACTTCATCGTGGGTGCTTCAACCGGGCTTCAAGGGGTCTCC 511
QY 266 AGGTACACCGGGCGGTGAAGCTGAGCGGTACAGCGCGCACCGACTTCCGCGCTCGAGAAC 207
DB 266 AGGTACACCGGGCGGTGAAGCTGAGCGGTACAGCGCGCACCGACTTCCGCGCTCGAGAAC 207
QY 512 GGTACCTCCGACCGCGAGGCTCGGCCACCGATCTCGGTGATGTGGAACCGAGTGC 571
DB 512 GGTACCTCCGACCGCGAGGCTCGGCCACCGATCTCGGTGATGTGGAACCGAGTGC 571
QY 206 AGGGCGGCTTGGTGGCTTGGAGGTCAATACCTGCGGGCGATGACCGCGCGAGCTGC 147
DB 206 AGGGCGGCTTGGTGGCTTGGAGGTCAATACCTGCGGGCGATGACCGCGCGAGCTGC 147
QY 572 TGCCAGCATCTGTCGGCGGTTCATCTGCTGATCGATCGATCGATCGATCGATCGATCGAT 621
DB 572 TGCCAGCATCTGTCGGCGGTTCATCTGCTGATCGATCGATCGATCGATCGATCGATCGAT 621
QY 146 TGCCGAGCGCGCGCGCGCGGTTCGCGGACTCTGCGAG 97
DB 146 TGCCGAGCGCGCGCGCGCGGTTCGCGGACTCTGCGAG 97

RESULT 9
US-09-410-551B-1
; Sequence 1, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
; FILE REFERENCE: 30062-20026.00
; CURRENT APPLICATION NUMBER: US/09/410,551B
; CURRENT FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 77536
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52275) ... (71465)
US-09-410-551B-1

Query Match          9.3%; Score 65; DB 4; Length 77536;
Best Local Similarity 44.7%; Pred. No. 0.00012;
Matches 294; Conservative 0; Mismatches 360; Indels 3; Gaps 1;

QY 41 GGGCGGTTCGCGACGCGCGCGCGATCGTGGCGGCGAGCGCTTCGCGGCGAGCGCTTCGAGAGTAGC 100
DB 41 GGGCGGTTCGCGACGCGCGCGCGATCGTGGCGGCGAGCGCTTCGCGGCGAGCGCTTCGAGAGTAGC 100
QY 101 GGTTCGAGGCGCGCGCGAGTGGGAGAGATCTCTTCGCGGGCGCTCGGTTCACAGAGGGCGCGA 160
DB 101 GGTTCGAGGCGCGCGAGTGGGAGAGATCTCTTCGCGGGCGCTCGGTTCACAGAGGGCGCGA 160
QY 71802 CCTGCTGAGACATCGAGGCGCGGGGACCGCGGACCGCTCATCTCCCGGTACGCCA 71861
DB 71802 CCTGCTGAGACATCGAGGCGCGGGGACCGCGGACCGCTCATCTCCCGGTACGCCA 71861
QY 161 TGTACTTCCACTTCGCTTCCAGGAGAGAGTGGCGCGCGGTGCTGCGCGGAGCAGACCC 220
DB 161 TGTACTTCCACTTCGCTTCCAGGAGAGAGTGGCGCGCGGTGCTGCGCGGAGCAGACCC 220
```

```

Db 71862 AGCGGCTGCCCTCCCTCGTCAATCAACGCGCTGTACCGGGCTCACCCCTGAGAGGGGGCGG 71921
QY 221 TGACGCTGGCGGTGCGCGAATCCGGCTCCAAAGCGCGAGGAACGTGTAGACCTCACCATGC 280
Db 71922 TGTGAGGACCGATGCGGACATCACCGGCTCGCGGATCTGGACAGGCTCAAGACG 71981
QY 281 TGTGCGCCACGGCATGCTGCACGATCCGATCTCGGGGGGGGACACGCGGCTCGCATCG 340
Db 71982 TGACCGACGACTTCTTTCGGGACGCGCTGCGGCTGCTCGCGGCGAAGCGGTGACGAGCGG 72041
QY 341 ACCAGGGGGGGTGGACTTCTCGACGCCAACCCTTTCGGCGAGTGGGGGACATCTGG 400
Db 72042 GCGAGGACTGCTGCAACCGGCTGCGGCTCGGCGACACGCGGAGATCTCGCTCAGCGACG 72101
QY 401 CCAGCTCTTGGCGGAGGACAGGAACGGGGGGAGTGTCTCGCACGCTGAACCCGAAAA 460
Db 72102 ACAGGCGGCGGCGTGTTCGCGACGCTGCTGTCGCGGCGCACGACTCGGTGCGACG 72161
QY 461 AGACGGCGCATTCATGCTGGCTGTCTTCAACGGGCTCGAGGCGGCTCCCGGGTCACT 520
Db 72162 TGTGCGGCTACTGCTCTACGCACTGCTCAGCCACCGCGGACAGCGGCGGCTCGCG 72221
QY 521 CGACCGCGGAGGCTCGGCGGCGGATCTCGATCGAAACCGGCGAGGAGCGGATCGGAGT 580
Db 72222 CGGCGCGGAGTGTGCGAACGCGGCTCGAGGAT---GCTCGTTTCTGCGCGTCA 72278
QY 581 TCGTCCGCGCTCCATGCTGACCTGATCGAAACCGGCGAGGAGCGGATCGGAGT 640
Db 72279 ACCAGATGGCGTACCGCGGCTGTGTGAGGACGTCGATGTGCGGGCGTGGCGATCC 72338
QY 641 CGGCGCGGCGGAGGCGGCGGAGGCTCGGAGGCGCTCGAGGCGGCTCGACGAGT 697
Db 72339 GTGCGGGCGACACGTCATCCGCTCTACTCGACGGCCAAACCGCGACCCCGAGGTGT 72395

RESULT 10
US-09-940-316B-1
; Sequence 1, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
; FILE REFERENCE: 30062-20026.11
; CURRENT APPLICATION NUMBER: US/09/940,316B
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/410,551
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 77536
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52275)...(71465)
US-09-940-316B-1

```

Query Match 9.3%; Score 65; DB 4; Length 77536;

Best Local Similarity 44.7%; Pred. No. 0.00012;

Matches 294; Conservative 0; Mismatches 360; Indels 3; Gaps 1;

```

QY 41 GGGCGCTCCGCAACGCGGAGGCGATCGTGGCGGCGAGCGCCCTCGGTCTTCGACGAGTACG 100
Db 71742 GGGACTTTCACACTGCGCGCGGCGCGAGCGGAGGACTTCTGTCGCCGAGGCGCGCGACG 71801
QY 101 GGTTCGAGGCGCGCCACGATGCGGAGATCTCTTCGCGGCGCTCGGTTCACAAAGGGCGCGA 160
Db 71802 CCGTCTGACGACATCGAGGCGCGGACCGCGGACCGACCTCATCCCGGGTACGCCA 71861
QY 161 TGTACTTCCACTTCGCTTCCAAAGGAAGTGGCGGCGGCGCTGTGGCGGAGGACGCC 220
Db 71862 AGCGGCTGCCCTCCCTCGTCAATCAACGCGCTGTACCGGGCTCACCCCTGAGAGGGGGCGG 71921
QY 221 TCCAGCTGCGGCTGCGGGAATCCGGCTCCAAAGCGCGAGAACTGTGTAGACCTCACCATGC 280
Db 71922 TGTGAGGCGCGATGCGCGACATCACCGGCTCGGCGGATCTGACAGCGTCAAGACG 71981
QY 281 TGTGCGCCACGGCATGTGTGACGATCCGATCTCTCGGGGCGGCGACGCGGCTCGCACTGG 340
Db 71982 TGACCGACGACTTCTTTCGGGCAACGCGCTGCGGCTGCTCGCGGAAAGGCTGACGAGCGG 72041
QY 341 ACCAGGGGGCGGTGAGCTTCTCCGACGCCAACCCTTTCGGCGAGTGGGGGCGACATCTGG 400
Db 72042 GCGAGGACTGCTGCAACCGGCTGCGGCTCGGCGACGACGCGGAGATCTCGCTCAGCGAG 72101
QY 401 CCAGCTCTTTCGCGGAGGCGACAGGAACCGGGGGAGGTGCTTCGCGACGCTGAACCCGAAAA 460
Db 72102 ACAGGCGGCGGCGTGTTCGCGACGCTGTGTCGCGGCGCACGACTCGGTGCGACG 72161
QY 461 AGACGGGCGACTTATCGTGGCTGCTTCAACGGGCTCGAGGCGGCTTCGCGGCTCACCT 520
Db 72162 TGTGCGGCTACTGCTCTACGCACTGCTCAGCACGCTCGACGACCCCGGAGCGGCGCTCG 72221
QY 521 CGACCGCGGAGGCTCGGCGGCGGATCTCGATCGAAACCGGCGAGGAGTGTGGAACGCTG 580
Db 72222 CGGCGCGGAGTGTGCGAACGCGGCTCGAGGAT---GCTCGTTTCTGCGCGTCA 72278
QY 581 TCGTCCGCGCTCCATGCTGATCGAAACCGGCGAGGAGCGGATCGGAGGAGTGG 640
Db 72279 ACCAGATGGCGTACCGCGGCTGTGTGAGGACGTCGATGTGCGGGCGTGGCGATCC 72338
QY 641 CGGCGGCGGCGGAGGCGGCGGAGGCTCGGAGGCGCTCGAGGCGGCTTCGAGCGAGT 697
Db 72339 GTGCGGGCGACAACTGATCCGCTCTACTCGACGGCCAAACCGCGACCCCGAGGTGT 72395

RESULT 11
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-08-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Query Match

9.0%; Score 62.8; DB 3; Length 4403765;

Best Local Similarity 49.2%; Pred. No. 0.00063; Matches 194; Conservative 0; Mismatches 197; Indels 3; Gaps 1;

QY 124 GAGATCCTCTCGCGGCGCTCGGTACCAAGGGCGCGATGTACTTCCACTTCGCTTCCAAG 183
Db 3647034 GGGATTGCGCGGATCGTCTCATCGCGGACCAACCCGGAGATCCGCCAGGTGTCGCC 3647093

QY 184 GAAGAGCTGGCGCGCGGTGTCGGCGGAGAGACCTCGACGTGCGGTGCGGGAATCC 243
Db 3647094 GAGCAATCGGGATTGATGAGTGGCGCGCGAGGTATCGCGGAGCAAGCTCGCGGG 3647153

QY 244 GGTCTCAAGCGCGAGAACTGGTGTGACCTCACCATCTGTGTGCGCCACCGCATCTGTCAC 303
Db 3647154 GTGCGGAGCTGCGAGACAGCGCTACGTCTGCGGATGTCGCGGACCGCATCAACGAC 3647213

QY 304 GATCGATCTCGCGGCGGCG---ACGCGGCTCGACTGGACACAGGGCGGTGGACTTC 360
Db 3647214 GCCCGGCGCTGCGCGCGCGGATATCGGGATCGGCCTTCCCGGACCGAGCTC 3647273

QY 361 TCGGAGCGCAACCGCTTGGCGAGTGGGGGACATCTGGCGCCAGCTCTGCGGGAGGCA 420
Db 3647274 GCGGTGCGAGACCGCGATGTCGCGCTGGCCAAAGACGACCTGACCGCCTGCTCGAGCTT 3647333

QY 421 CAGGAACGGGGAGGTGCTTCGCGACGTGAACCCGAAAGACCGGACTTCATCGTC 480
Db 3647334 GGGGACTGGCGAGCGCGGAGTGTATTCGCGAGAACTACCGCATGTCCATCGCC 3647393

QY 481 GCGTGTTCACCGGCTCGAGGCGGTCTCCCGGG 514
Db 3647394 GTCAACGCGCGCGGTCTGATCGCGCGGGCG 3647427

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; ORIGIN INFORMATION: H37Rv
US-09-103-840A-1

Query Match 9.0%; Score 62.8; DB 3; Length 4411529;
Best Local Similarity 49.2%; Pred. No. 0.00063;
Matches 194; Conservative 0; Mismatches 197; Indels 3; Gaps 1;

QY 124 GAGATCCTCTCGCGGCGCTCGGTACCAAGGGCGCGATGTACTTCCACTTCGCTTCCAAG 183
Db 3652184 GGGATTGCGCGGATCGTCTCATCTCACCGGACCAACCCGGAGATCCCGAGGTGTGCGC 3652243

QY 184 GAAGAGCTGGCGCGCGGTGTCGGCGGAGACACCTCGACGTGCGGTGCGGGAATCC 243
Db 3652244 GAGCAATCGGGATTGATGAGTGGCGCGCGAGGTATCGCGGAGCAAGCTCGCGGG 3652303

QY 244 GGTCTCAAGCGCGAGAACTGGTGTGACCTCACCATCTGTGTCGCCACCGCATCTGTCAC 303
Db 3652304 GTGCGGAGCTGCGAGACGACGCTACGTCTGCGGATGTCGCGGACCGCATCAACGAC 3652363

QY 304 GATCCGATCTCGCGGCGGCG---ACGCGGCTCGACTGGACACAGGGCGGTGGACTTC 360

Db 3652364 GCCCGGCGCTGCGCGCGCGCGATATCGGGATCGCATCGGCTTGCAGAACCGAGCTC 3652423

QY 361 TCCGAGCGCAACCCGTTGCGGAGTGGGCGGACATCTGCGCCAGCTCTTGGCGGAGGCA 420
Db 3652424 GCGGTGAGAGCGCGCGATGTGCGGTGGCCAAAGACGACCTGCACCGCTGTCTCGAGCTT 3652483

QY 421 CAGGAACGGGGAGGTGCTTCGCGACGTGAACCCGAAAGACCGGCGACTTCATCGTC 480
Db 3652484 GGGGACCTGGGCGAGCGGCGAGTGTGATGTAATCCGCGAGAACTACGGCATGTCCATCGCC 3652543

QY 481 GCGTGTCTTCAACCGGCTCGAGGCGGTCTCCCGGG 514
Db 3652544 GTCAACGCGCGCGGTGCTGATCGCGCGGGCG 3652577

RESULT 13
US-09-902-540-4889
; Sequence 4889, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4889
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-4889

Query Match 8.9%; Score 62; DB 4; Length 630;
Best Local Similarity 48.3%; Pred. No. 0.00026;
Matches 204; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 34 CAGGAACGGGCGTCCGCGCGCGGCGATCGTGGGCGAGCGCGCTCGGTCTTCGAC 93
Db 37 CAGGAGCGGCGCGGTACGCGCGCGCGCTGTGATGGAGCGGCCATTTGGCGGCTCTCC 96

QY 94 GAGTACGGGTTCGAGCGCGCGCGCGAGTCTCTCGGGGCGCTCGGTCAACGAG 153
Db 97 GAGTGGGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 156

QY 154 GCGCGATGTACTTTCATCTTCCAAAGGAGAGCTGGCGCGCGCGCGCGCGCGCGCG 213
Db 157 GCGCGCTGTGACGACCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 216

QY 214 CAGACCTGCGCGGTGCGGAAATCCCGCTCCAAAGCGCGAGGAATCGGTAGA--C 270
Db 217 GTGGGCGCGCGCGGTGAGCGCGGTCTCCGCAAGCGCGCGCGCGCGCGCGCGGT 276

QY 271 CTCAACATGCTGTGCG 330
Db 277 CG 336

QY 331 CTGCGACTGAGACAGGGCGGTGAGCTTCTCGAGCGCGCGCGCGCGCGCGCGCGCG 390
Db 337 GCGCGGTGCGCGCTCTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 396

QY 391 GACATCTGCG 450
Db 397 CCGCTGGAGCG 456

QY 451 AA 452
Db 457 GA 458

RESULT 14
US-09-902-540-1241
; Sequence 1241, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1241
; LENGTH: 28493
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(28493)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1241

Query Match 8.9%; Score 62; DB 4; Length 28493;
Best Local Similarity 48.3%; Pred. No. 0.00045;
Matches 204; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 34 CAGGAACGGGCGGCTCCGACGCGGAGGCGATCTGTCGGGGAGCGCCCTCGGTCTTGGAC 93
DB 6110 CAGGAGCGCAGCGCGTCAAGCGGAGCTCTGATGGAGCGGCGCATTTGGCGGTCTCC 6169

QY 94 GAGTACGGTTGAGCGCCGCAAGTGGCAGAGATCTCTCGGGGCTCTCGGTCAACAG 153
DB 6170 GAGCTGGGTGGGCGGCGCGACCATGACTGTGATTGCGAGCGCGCGGTGTGCCGG 6229

QY 154 GCGCGATGCTACTTCCACTTCCTTCCAGAGAGAGCTGCGCGCGGCGGTCTGTCGCGAG 213
DB 6230 GCGCGCTGTGAGCACCATTCCTCCACCCCGCGCGACCTGTGGCGCGCGCGGTGTGTC 6289

QY 214 CAGACCTGACGTGCGGCTCGGGAATCCGGCTCCAGGCGCAGGAATGTTAGATA---C 270
DB 6290 GTGGGSCACAGCAGGTGAGCGGTGCTCCGAGGCGCGCATTTGCGCGCGATGCT 6349

QY 271 CTCACCATCTGTCGCGCCACGGCATCTGTCAGGATCCGATCTGCGGGCGGGCAGCGG 330
DB 6350 CGCCGACAGGAGCCATCTCCACATGCTGCGCGCGCATCTACTCAGCCCTGTCTTACC 6409

QY 331 CTCGCACTGACACAGGGGCGGTGGACTTCTCCGACGCCAACCCGTTGCGCGAGTGGGCG 390
DB 6410 GCGGGGTGACGTCTGGGTGGCGCGGTGCGGAGCGCGAGCTGCGCGCGCATTTGGCC 6469

QY 391 GACATCTGCGCCAGCTCTGCGGAGGACAGGAACGCGGGGAGGTGCTTCCGACAGTG 450
DB 6470 CCGTGGAGCGCGGTGGGGCGGAGTTCCACCGGCTGACGCTGACGTTGCTCGGGTG 6529

QY 451 AA 452
DB 6530 GA 6531

RESULT 15
US-09-773-816-1
; Sequence 1, Application US/09773816
; Patent No. 6340774
; GENERAL INFORMATION:
; APPLICANT: Stanford University
; APPLICANT: Khosla, Chaitan
; TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR
; TITLE OF INVENTION: ANTAGONISTS

FILE REFERENCE: 28600-20210.00
; CURRENT APPLICATION NUMBER: US/09/773,816
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/243,458
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/179,305
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23673
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(23623)
; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1

Query Match 8.6%; Score 60.2; DB 3; Length 23673;
Best Local Similarity 45.0%; Pred. No. 0.0011;
Matches 313; Conservative 0; Mismatches 373; Indels 9; Gaps 2;

QY 14 ACGAACGGGTGGCACTGCGACGAGGAACGGGGCGGTCCGCA CGCGCAGCGCATCTGTCGGG 73
DB 15958 ACACGCCGTGGCGCTGTACGTGCACGCGCTGGGCGTCAACCGGTGTGCGACGGTCGAGT 16017

QY 74 CAGCGCGCTCGGTCTTCAGCAGATACGGGTTTCAGGCGCGCACATGTCGAGAGATCTCTT 133
DB 16018 TCGACGCTCCGCGCGACCGACGAGGACATGACCGCGAGGAGTGTGCGGCGCAGCATCC 16077

QY 134 CGCGCGCTCGGTCTACCAAGCGCGATGTACTTCCACTTCGCTTCCAAAGGAAGAGCTGG 193
DB 16078 GACGGGTGAGCGGACGAGCGTGCAGATCAGGAGATGTCTTCGGGACCCCGTGCATCG 16137

QY 194 CCGCGCGGTGTGTCGCGAGCAGACCTTCGACGTGGCGGTGCGGGAATCCGGCTCCAAAG 253
DB 16138 ACACGCCGCGCAGCGCTCCGCTACCGCTCGCGCGGTGTGCTGGCGCGGCGAGCGG 16197

QY 254 CGCAGGAAGTGTAGACCTCACCATGTCTGGTCCGCCACGCGCATGTCGACGATCCGATCC 313
DB 16198 CGCAGTGTACGCGCGGTGCGCGCGCGCGCTCAACGTCGCGCTGTGTCGACGCGGCGCA 16257

QY 314 TCGCGCGCGGCA CGCGGTCTGCATCTGACACGAGGGGCGGTGAGATCTTCTCCGACGCCAAC 373
DB 16258 ACCTGCTGTGAGAGTGTGCGCGCGGAGGTACAGCGCTGGCGCGCGCATCTATCTCTCGACA 16317

QY 374 CGTTCCGCGAGTGGGCGACATCTGCGCCCGAGCTCTGCGCGGAGGCA CAGGAACGCGGGG 433
DB 16318 CTTACGTTTCGAGAGCGCTACCGCGTCCGCGCGGTGTTGCAGAAACACCGCGCGGCGAGA 16377

QY 434 AGGTGCTTCCGCACTGAAACCGGAAGAGACCG-----GGACTTCAATCTGTCGGCTGCT 487
DB 16378 TCGCGCTCATGACCCCGCGCGCGAGCACCGCGCTCGCGGAGATGTTTCAGAGGAATGCG 16437

QY 488 TCACCGGCTCCAGCGGTCTCCCGGGTCACTCTCCGACCGCGAGGACCTCGCGCACCGGGA 547
DB 16438 TCGACATCAGCAGGTGTCACCGCAGATCGCGGATGATGCGCGGAATGAGCGTGGGT 16497

QY 548 TCTCGGTGATGTGAAACCAAGTGTGCGCCAGCA---TCGTGCGGCGGTCCATGTCGACCT 604
DB 16498 ACGCCGTGAGGCGGAGCATCCGCTCTCTCGGCAAGCTGCTGCGGACCGCAAGCTCACCG 16557

QY 605 GGATCGAAACCGCGGAGGAGCGGATCGGGAAGTTCGCGCGCGCGCGCGCGCGCGCGGAGG 664
DB 16558 GCGCAAGCGCGCGCACTGTGGGCGGACGCGCGCGGTGCTCTGCTGCTGCTGCTGCTGCTG 16617

QY 665 CTGCGGAGGCTCCGAGCGCGCTCCGACGAGTAG 699
DB 16618 GCGCGAGCTGCGCGACGCGGTGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 16652

Search completed: March 6, 2005, 22:16:10

Mon Mar 7 10:28:17 2005

us-10-049-710a-1.rni

Page 9

Job time : 242 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:57:05 ; Search time 630 Seconds
(without alignments)
6583.744 Million cell updates/sec

Title: US-10-049-710A-1

Perfect score: 699

Sequence: 1 atggcagtcgcacgaacg.....agggcgctccgacgagtag 699

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5401638 seqs, 2966923429 residues

Total number of hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	699	100.0	699	15 US-10-049-710A-1	Sequence 1, Appli
2	229	32.8	4346	15 US-10-017-471A-19	Sequence 19, Appl
3	157.4	22.5	702	15 US-10-156-761-3692	Sequence 3692, Ap
4	157.4	22.5	9025608	15 US-10-156-761-1	Sequence 1, Appli
5	112.4	16.1	9025608	15 US-10-156-761-1	Sequence 1, Appli
6	112.2	16.1	654	15 US-10-156-761-1	Sequence 3689, Ap
7	93.4	13.4	654	15 US-10-156-761-3689	Sequence 18, Appl
8	85.8	12.3	9521	17 US-10-168-663-18	Sequence 2257, Ap
9	85.8	12.3	9521	17 US-10-168-663-18	Sequence 19, Appl
10	78.6	11.2	851	17 US-10-156-761-3690	Sequence 3690, Ap
11	78.2	11.2	4257	9 US-09-825-288A-1	Sequence 1, Appli

12	69.6	10.0	18438	15 US-10-156-761-2886	Sequence 2886, Ap
13	69.6	10.0	125746	15 US-10-156-761-15102	Sequence 15102, A
14	68.4	9.8	3173	18 US-10-437-963-77080	Sequence 77080, A
15	68.2	9.8	1324	17 US-10-282-132A-25841	Sequence 25841, A
16	68.2	9.8	32329	17 US-10-374-903A-1	Sequence 1, Appli
17	67.2	9.6	1041	9 US-09-815-242-7885	Sequence 7885, Ap
18	67.2	9.6	1041	17 US-10-282-122A-30427	Sequence 30427, A
19	66.4	9.5	603	15 US-10-156-761-4514	Sequence 4514, Ap
20	65	9.3	657	15 US-10-156-761-5997	Sequence 5997, Ap
21	65	9.3	77536	10 US-09-940-316B-1	Sequence 1, Appli
22	64	9.2	9521	17 US-10-168-663-18	Sequence 18, Appl
23	64	9.2	9521	17 US-10-168-663-18	Sequence 19, Appl
24	63.4	9.1	5862	16 US-10-132-134-15	Sequence 15, Appl
25	63.4	9.1	52101	16 US-10-132-134-1	Sequence 1, Appli
26	63.2	9.0	681	15 US-10-156-761-6081	Sequence 6081, Ap
27	63	9.0	591	15 US-10-156-761-4766	Sequence 4766, Ap
28	62.8	9.0	618	15 US-10-156-761-2255	Sequence 2255, Ap
29	62.6	9.0	582	17 US-10-282-122A-11537	Sequence 11537, A
30	62.6	9.0	633	15 US-10-156-761-558	Sequence 558, App
31	62.4	8.9	819	17 US-10-282-122A-11958	Sequence 11958, A
32	62.4	8.9	1581	11 US-09-758-759-122	Sequence 122, App
33	62.4	8.9	109519	11 US-09-758-759-1	Sequence 1, Appli
34	61.4	8.8	651	15 US-10-156-761-744	Sequence 744, App
35	60.8	8.7	3516	18 US-10-437-963-74025	Sequence 74025, A
36	60.8	8.7	5877	14 US-10-152-886-94	Sequence 94, Appl
37	60.6	8.7	381	18 US-10-021-323-17015	Sequence 17015, A
38	60.2	8.6	1395	15 US-10-156-761-2088	Sequence 2088, Ap
39	60.2	8.6	1836	15 US-10-156-761-2659	Sequence 2659, Ap
40	60.2	8.6	27541	18 US-10-203-295-2	Sequence 2, Appli
41	60.2	8.6	125401	18 US-10-203-295-35	Sequence 35, Appl
42	60	8.6	1893	15 US-10-156-761-5675	Sequence 5675, Ap
43	60	8.6	27541	18 US-10-203-295-2	Sequence 2, Appli
44	60	8.6	125401	18 US-10-203-295-35	Sequence 35, Appl
45	59.8	8.6	573	15 US-10-156-761-4685	Sequence 4685, Ap

ALIGNMENTS

RESULT 1

US-10-049-710A-1
; Sequence 1, Application US/10049710A
; Publication No. US20030126648A1
; GENERAL INFORMATION:
; APPLICANT: Shinmyo, Atsuhiko
; APPLICANT: Kato, Kou
; APPLICANT: Yamada, Yasuhiro
; APPLICANT: Nihira, Takuya
; APPLICANT: Shindo, Takuya
; TITLE OF INVENTION: METHOD FOR INDUCTION OF GENE EXPRESSION IN PLANT AND PLANT TREAT
; FILE REFERENCE: 5405/18
; CURRENT APPLICATION NUMBER: US/10/049,710A
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: PCT/JP01/05096
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: JP 2000-180466
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

LENGTH: 699

TYPE: DNA

ORGANISM: Streptomyces virginiae

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(699)

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: Okamoto, S., Nakamura, K., Nihira, T. and Yamada, Y.

TITLE: Virginiae butanolide binding protein from Streptomyces virginiae.

TITLE: Evidence that VbrA is not the virginiae butanolide binding protein and re-

JOURNAL: Journal of Biological Chemistry

VOLUME: 270

ISSUE: 20

PAGES: 12319-12326

DATE: 1995-05-19

DATABASE ACCESSION NUMBER: D32251

DATABASE ENTRY DATE: 1994-07-19

RELEVANT RESIDUES: (1)..(699)

PUBLICATION INFORMATION:

AUTHORS: Okamoto, S., Nakamura, K., Nihira, T. and Yamada, Y.

TITLE: Virginiae butanolide binding protein from Streptomyces virginiae.

TITLE: Evidence that VbrA is the true virginiae butanolide binding protein and re-

TITLE: Identification of the true binding protein

JOURNAL: Journal of Biological Chemistry

VOLUME: 270

ISSUE: 20

PAGES: 12319-12326

DATE: 1995-05-19

DATABASE ACCESSION NUMBER: D32251

DATABASE ENTRY DATE: 1994-07-19

US-10-049-710A-1

Query Match 100.0%; Score 699; DB 15; Length 699;

Best Local Similarity 100.0%; Pred. No. 6e-172;

Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCAGTCCGACACGAAACGGGTGGCAGTGCAGACGAGAACGGCGCTCGCACGCGGCAG 60
Db 1 ATGGCAGTCCGACACGAAACGGGTGGCAGTGCAGACGAGAACGGCGCTCGCACGCGGCAG 60

QY 61 GCGATCGTCGGGCGACGCCCTCGGTCTTCGACGAGTACGGGTTCGAGGCGGCCACAGTG 120
Db 61 GCGATCGTCGGGCGACGCCCTCGGTCTTCGACGAGTACGGGTTCGAGGCGGCCACAGTG 120

QY 121 GCAGAGATCCTCTCGGGGCGCTCGGTACACAGAGGCGCGATGACTTCCACTTCGTTCC 180
Db 121 GCAGAGATCCTCTCGGGGCGCTCGGTACACAGAGGCGCGATGACTTCCACTTCGTTCC 180

QY 181 AAGGAAGAGTGGCGCGCGCGTGTGGCGAGCAGACCTGCAGTGGCGGTGCGGAA 240
Db 181 AAGGAAGAGTGGCGCGCGCGTGTGGCGAGCAGACCTGCAGTGGCGGTGCGGAA 240

QY 241 TCGGCTCAAGCGCGAGAACTGGTAGACTTCAACATGCTGGTTCGCCACCGGCATGCTG 300
Db 241 TCGGCTCAAGCGCGAGAACTGGTAGACTTCAACATGCTGGTTCGCCACCGGCATGCTG 300

QY 301 CAGGATCCGATCTCGGGGCGGCGACGCGCTCGCACTGGAACGAGGCGCGTGGACTTC 360
Db 301 CAGGATCCGATCTCGGGGCGGCGACGCGCTCGCACTGGAACGAGGCGCGTGGACTTC 360

QY 361 TCCGAGCGCAACCCGTTCCGCGAGTGGGCGACATCTGCGCCGAGCTCTCTGGCGGAGCA 420
Db 361 TCCGAGCGCAACCCGTTCCGCGAGTGGGCGACATCTGCGCCGAGCTCTCTGGCGGAGCA 420

QY 421 CAGGAAACGGGGGAGGTGTTTCGCACTGTAACCCGAAAGAACCGGCGACTTCATCGTC 480
Db 421 CAGGAAACGGGGGAGGTGTTTCGCACTGTAACCCGAAAGAACCGGCGACTTCATCGTC 480

QY 481 GGTGCTTCAACGGGCTCAGGGGTCTCCGGGTCACTCGGACCGCGGCGAGCTCGGC 540
Db 481 GGTGCTTCAACGGGCTCAGGGGTCTCCGGGTCACTCGGACCGCGGCGAGCTCGGC 540

QY 541 CACCGGATCTCGGTGATGTGGAACACGAGTGTGCCAGCATCTGTGCGGCGCTCCATGCTG 600
Db 541 CACCGGATCTCGGTGATGTGGAACACGAGTGTGCCAGCATCTGTGCGGCGCTCCATGCTG 600

QY 601 ACTTGATCGAAACCGCGAGAGCGGATCGGAAGGTTCGGGCGGCGCGAGCGCGCC 660
Db 601 ACTTGATCGAAACCGCGAGAGCGGATCGGAAGGTTCGGGCGGCGCGAGCGCGCC 660

QY 661 GAGGCTCGGAGGCTCCGAGGCGCGCTCCGACGAGTAG 699
Db 661 GAGGCTCGGAGGCTCCGAGGCGCGCTCCGACGAGTAG 699
```

RESULT 2

US-10-017-471A-19

Sequence 19, Application US/10017471A

Publication No. US2003012464A1

GENERAL INFORMATION:

APPLICANT: Takano, Eriko

APPLICANT: Bibb, Mervyn

TITLE OF INVENTION: Antibiotic Production

FILE REFERENCE: 0380-P02329US1

CURRENT APPLICATION NUMBER: US/10/017,471A

CURRENT FILING DATE: 2001-10-23

PRIOR APPLICATION NUMBER: US 60/242,561

PRIOR FILING DATE: 2000-10-23

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 19

LENGTH: 4346

TYPE: DNA

ORGANISM: Streptomyces coelicolor

US-10-017-471A-19

Query Match 32.8%; Score 229; DB 15; Length 4346;

Best Local Similarity 60.8%; Pred. No. 5.2e-50;

Matches 373; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

```
QY 34 CAGGAACGGGCGTCCGACGCGCGGCGGATCGTGGGCGAGCGCCTCGGTCTTCGAC 93
Db 2270 CAGGAACGGGCGGATCCGACGCGCGGCGGATCTTGGAGCGCGCGCGGCGGATCTTCGAG 2329

QY 94 CAGTACGGGTTCAGGCGCGCCACAGTGGCAGAGATCTCTCGGGGCGCTCGGTTCACCAAG 153
Db 2330 AAGCAGGCTACCAAGTGCACAGATCAGGAGATCTCAAGTGGCGGGGTGACCAAG 2389

QY 154 GCGCGATGTACTTCCACTTCGCTTCCAGAGAGAGTGGCGCGCGCGGCGGTGTCGCCGAG 213
Db 2390 GAGAGCCCTCTACTTCCACTTCAGTCCAGGAGAACTGGCGCTGGGCGGTCTTCGACGCC 2449

QY 214 CAGACCTTCGACGTGGCGGTGCGGAATCGCGCTCCAAGGCGCAGAACTGGTAGACCTC 273
Db 2450 CAGGAACCAACACAGCGGCTTCGAGGCAACCCCTCCGGCTGCAAGAACTCATCGACATG 2509

QY 274 ACCATGCTGTCGCCCGACGGCATGTGCAAGATCCGATCTCTCGGGCGGCGACGCGGCTC 333
Db 2510 GGCATGTTGTTCTGTCACTGCGTTCGCGACGAACTGCGGCGCGCGCGCGCGCTC 2569

QY 334 GCACCTGGACAGGCGGCGGTGCACTTCTCCAGGCGCAACCCGTTCCGCGAGTGGGCGGAC 393
Db 2570 TCCATGGACAGCAGGCGCACCGGTCTCGATTCGCGAGGACCTTCGCTCGCTGGCACGAG 2629

QY 394 ATCTCGCGCCAGCTCTCTGGCGGAGGACAGGAAACGGGGGAGGTGTTCCGCACTGAAAC 453
Db 2630 ACATCTCTGAAGTGTGTAACAGGCGCAGAGAGAGCGTGTGCTGCCCATGTGGTC 2689

QY 454 CCGAAAAAGACCGCGCATTTTCATCTGTCGCTGCTTCAACGCGGCTCCAGCGGTCTCCCGG 513
Db 2690 ACCACCGCATCTCGCGCATCTCTAATGCGGACGTTTCGCGGATACAGGTCTGTGTCGCG 2749

QY 514 GTCACCTCGACCGCGCAGGACCTCGGCGCACCGGATCTCGGTGATGTGGAACCACTGCTG 573
Db 2750 ACGGTGACGACATACAGGACCTCGAAACCCCTACGCGCTCTCTCGAGAGACATCTCTG 2809

QY 574 CCGAGCATCTGTCGCGGTCTCACTGTGATCGATCGAAACCGGCGGAGGCGGATCGGG 633
Db 2810 CCGCGCATCTGCGGTCTCCCTCGGTGCTGGCGCGCTCGATCTCTCCGAGAGCGCGGAGCA 2869

QY 634 AAGGTGCGGCGG 646
Db 2870 CGCCTCGGCGCG 2882
```

RESULT 3

US-10-156-761-3692
; Sequence 3692, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3692
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(702)
US-10-156-761-3692

Query Match 22.5%; Score 157.4; DB 15; Length 702;
Best Local Similarity 55.8%; Pred. No. 2.2e-31;
Matches 371; Conservative 0; Mismatches 276; Indels 18; Gaps 3;

QY 28 GTGCGCAGAGAAAGCGGCGCTCCGACGCGGCGAGCGATCGTGGCGCAGCGCGCTCGGTC 87
DB 4 GCGCGCAGAGAGCGGCGCTCCGACGCGGCGAGCGATCTGTGTCGCGCGCGCGAGGTG 63

QY 88 TTGACGAGTACGGGTTCGAGGCGCGCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTC 147
DB 64 TTGACGAGTGGGATACGAGGCGGCAACCATCTCCGACGCTGCTGAAGCGCTCGGCGGTC 123

QY 148 ACCAAGGCGCGATGACTTCCACTTCGCTTCCAGGAGAGAGTGGCGCGCGCGTGGCTG 207
DB 124 ACCAAGGCGCGCTCTACTTCCACTTCAGTCCAGAGAGAGTGGCGCGCGCGTGGCTG 183

QY 208 GCGGAGCAGACCTGACAGCGTGGCG---GTGCGCGAATCCGGCTCCAGGCGCAGGAACTG 264
DB 184 GCGGAGCAGTCCGCTCCCTCCGCGCTCCCGAGCAGGAGCTGAAGCTCCAGCAGTGC 243

QY 265 GTAGACCTCACCAGTGTGTCGCGCACCGGATCTGCA-----CGATCCGATCTG 315
DB 244 CTGGACGAGCGCTGCTGCTCGCCATCTGCTCAGGGAAGGCAACCGCGATCCGATCTG 303

QY 316 CGGCGGCGCAGCGGCTCGACTGACACCGAGGCGG---GTGCGACTTCTCCGAGCGC 369
DB 304 CAGGCGAGTGTGGGCTGACCGTGGACCGAGGCTCCCGCAGGAGCATTCTCAACCGCGG 363

QY 370 AACCCCTTCGCGAGTGGGCGACATCTGCGCCAGCTCTGCGGAGGCGACAGGAACTG 429
DB 364 GTCCCATGACAGGCTGGACCGAGCACACGAGTCTCTTTCGAAGAGGCGAGGCGAAG 423

QY 430 GGGGAGGTGTTCCGACGTGAACCCGAAAGACCGGCGACTTCATGTCGCTGCTTC 489
DB 424 GCGGAGATCTGCGCCACGCGGATGTGAAGCGCTCGCAAGCTGTGTCGCGCGGCTTC 483

QY 490 ACCGGCTCAGGCGCTCCCGGCTCCTCCGACCGCAGGAGCTCCGCGCACCGGATC 549
DB 484 ACCGGCTCAGGCTCTCTCAGGATCATGACCGGCGCGGAGCTCCGCGAGCGGCTG 543

QY 550 TCGGTGATGTGAACACAGTGTGCGCAGCATCTGCGCGCGTCCACTGTGACTGATC 609
DB 544 GCGGAGCTTACCGCATCTGATCGCTCTTCGCGCATGCGGGGATCTGTCGCGCTG 603

QY 610 GAAACCGCGCAGAGCGGATCGGAAGTTCGCGGCGCGCGCGCGCGCGCGCGCTGCG 669
DB 604 GACTTCTCCCGAGCGCGGCTCGCGGTGTACGAAGCGCGCATGAAGCAGCGGAGTGC 663

QY 670 GAGGC 674
DB 664 GCGGC 668

RESULT 4
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 22.5%; Score 157.4; DB 15; Length 9025608;
Best Local Similarity 55.8%; Pred. No. 1.2e-31;
Matches 371; Conservative 0; Mismatches 276; Indels 18; Gaps 3;

QY 28 GTGCGCAGAGAAAGCGGCGCTCCGACGCGGCGAGCGATCGTGGCGCAGCGCGCTCGGTC 87
DB 4585273 GCGCGCAGAGAGCGGCGCTCCGACGCGGCGAGCGATCTGTGTCGCGCGCGAGGTG 4585214

QY 88 TTGACGAGTACGGGTTCGAGGCGCGCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTC 147
DB 4585213 TTGACGAGTGGGATACGAGGCGGCAACCATCTCCGAGCTGCTGAAGCGCTCGGCGGTC 4585154

QY 148 ACCAAGGCGCGATGACTTCCACTTCGCTTCCAGGAGAGAGTGGCGCGCGCGTGGCTG 207
DB 4585153 ACCAAGGCGCGCTCTACTTCCACTTCAGTCCAGCAGAGAGTGGCGCGCGCGTGGCTG 4585094

QY 208 GCGGAGCAGACCTGACAGCTGGCG---GTGCGCGAATCCGGCTCCAGGCGCAGGAACTG 264
DB 4585093 GCGGAGCAGTCCGCTCCCTCCGCGCTCCCGAGCAGGAGCTGAAGCTCCAGCAGTGC 4585034

QY 265 GTAGACCTCACCAGTGTGTCGCGCACCGGATCTGCA-----CGATCCGATCTG 315
DB 4585033 CTGGACGAGCGCTGCTGCTCGCCATCTGCTCAGGGAAGGCAACCGCGATCCGATCTG 4584974

QY 316 CGGCGGCGCAGCGGCTCGACTGACACCGAGGCGG---GTGCGACTTCTCCGAGCGC 369
DB 4584973 CAGGCGAGTGTGGGCTGACCGTGGACCGAGGCTGCCCGAGGAGCATTCTCAACCGCGG 4584914

QY 370 AACCCCTTCGCGAGTGGGCGACATCTGCGCGCGACTCTGCGCGGAGGCACAGGAACTG 429
DB 4584913 GTCCGATGACGCGCTGGACCGAGCAGCAGCAGTCTCTTTCGAAGAGGCGCGGCGAAG 4584854

QY 430 GGGGAGGTGCTTCCGACGTGAACCCGAAAGACCGCGGAGCTTCTATCTGCTGCGCTGCTTC 489

Db 4584853 GCGGAGATCTGCCCCACGCCGATGTGGAAGCGCTCGCAAGCTGTGTGTTGGCGCGTTTC 4584794
QY 490 ACCGGGCTCAGCGGCTCTCCGGGTACCTCGACCGCCAGGACTCGGCCACCGGATC 549
Db 4584793 ACCGGGCTCAGGTCCTCTCGAGGATCATGACCGGGCGCGGAGCTTGGCGAGCGGGTG 4584734
QY 550 TCGGTGATGTGGAACACAGTGTGCCAGCATCTGCGGGCTGCCATGTGACCTGGATC 609
Db 4584733 GCGGACCTTACCGCATCTGATGCGCTTTCGCCATGCCGGGATCTGTTCGCCCTG 4584674
QY 610 GAAACCGGAGAGCGGATCGGGAAGGTTCGGCGCGCGCCGAGCGCCCGAGGCTGG 669
Db 4584673 GACTTCTCCCGAGCGGGCTCGCGGTGTACGAAGCCGCATGAAGCAGCGGAGTCG 4584614
QY 670 GAGGC 674
Db 4584613 GCGGC 4584609

RESULT 5

US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 16.1%; Score 112.4; DB 15; Length 9025608;
Best Local Similarity 50.1%; Pred. No. 5.9e-20; Mismatches 326; Indels 15; Gaps 2;
Matches 343; Conservative 0;
QY 4 GCAGTGCACACAAACGGGTGGCAGTGCACAGAAACGGGCGCTCGCGACCGCGGAGCGG 63
Db 4581408 GCCTTCGCGAGGAGTCAACGGGTGACAAACAGGACGGCGCGCCGACCCGCCACGCC 4581467
QY 64 ATCGTGGGGCAGCGGCTCGCTTCGACGAGTACGGGTTGAGGCGCGCACAGTGGCA 123
Db 4581468 CTCATCGCTCCCGCGCCACCGCTTCGAAACGGCAGGGCTACACACAGGCGAGGCTGGCC 4581527
QY 124 GAGATCTCTCGCGGGCTCGGTCAACAGGGCGGATGACTTCACTTCGCTTCCAAG 183
Db 4581528 GACATAGCGCTTCGCGCGGTGTACGCCCGGCGACTGACTTCACTTCGAGAGCAAG 4581587
QY 184 GAGAGTGGCCCGCGCGGTGCTGGCGAGCAGACCTGCACTGCGGCGTTCGCGAATCC 243
Db 4581588 GCAGAGTGGCCAGGGCGGTGGAGGGCGCGCGGGGTGAGCTTGGCCCGCGGCGCTGG 4581647
QY 244 GGCTCCAGGCGCAGGAACTGGTAGACCT-----CACCATGCTGGTTCGCCAC 291

Db 4581648 CTGCCCCAGCGCGCGGCAAGAACCGCTGCAACCGGCTGACAGAACACATGTCACGCCCTG 4581707
QY 292 GGCATGCTGCACGATCCGATCTCGCGCGCGGACACGCGGCTCGCACTGACACAGCGGGCG 351
Db 4581708 GCGGAGCAGTTCGCGGGGAGTGTGTGTCGCCCGCGGGGCTTCGGCTGAACTGCCAATCG 4581767
QY 352 GTGACTTCTCCGACGCCAACCCGTTCCGGGAGTGGGCGCATCTGCGGCCAGCTTCCTG 411
Db 4581768 GCGGCGCGCGCGCTGAATCTGTCTCGGGAATGACAGACCTGCTGTGAGCAGCTGTCTC 4581827
QY 412 GCGGAGGACAGGAACGCGGGAGGTCTTCGCGACGTTGAACCCGAAAAAGACCGCGAC 471
Db 4581828 GCGGAGGCGCGCGAGGAGGCGTGTGTCGCGCGCTTCGTCGCGCAGACAGGTGACG 4581887
QY 472 TTCATGCTGGCTGTCTTCAACGGGCTCCAGGCGGTCTCCCGGGTCACTTCGACCGCCAG 531
Db 4581888 GCGGTGTGGCGCGACCAACCGGTTTCGAACTGCTCGCGCGCGGACCCCGAGTGGCT- 4581946
QY 532 GACTTCGCGCCACCGGATCTCGGTGATGTGGAACCAACGTCGTGTCGCCAGCATTCGTGCGGCG 591
Db 4581947 --CTCCGGCAGTCTGCTGCGCGCTTCTGCGGGTACTGTGCTGCGCGCGCGCGCG 4582004
QY 592 TCCATGCTGACTGATCGAAACCGCGGAGGAGCGGATCGGGAAGTTCGCGCGCGCGCC 651
Db 4582005 GCGGCGCTGACCGCGGTGACCCGACGAGAGTCCCGCAGCGCGGAGACCGGACC 4582064
QY 652 GAGGCGCGCGAGCTGCGGAGGCC 675
Db 4582065 CCGGCGACCAACCGCGGATGAGCC 4582088

RESULT 6

US-10-156-761-3689
; Sequence 3689, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3689
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(654)
US-10-156-761-3689

Query Match 16.1%; Score 112.2; DB 15; Length 654;
Best Local Similarity 50.9%; Pred. No. 1.2e-19; Mismatches 303; Indels 15; Gaps 2;
Matches 329; Conservative 0;
QY 33 ACAGGAACGGCGCTCGCGACCGCGGAGGATCGTGGCGGAGCGCGCTTCGGTCTTCGA 92
Db 9 ACAGGAACGGCGCGCGCGCACCGCCATCGCTTCATCGCTCGCGCGCCACCGCTTCGA 68
QY 93 CGAGTAGCGGTTGAGGCGCGCCACAGTGGCAGAGATCTCTTCGCGGGCTTCGGTCAACAA 152
Db 69 ACGGCAGGGCTACACACAGGCGAGGCTGGCCGACATCAGCGCTGCGCGCGGTGTGAGCCC 128

QY 153 GGGCGGATGTACTTCCACTTCCCTTCCAAAGGAAGAGTGGCCCGCGGGGTGCTGSCCGA 212
Db 129 CGCGCACTGCATCTTCCACTTCAGAGCAAGGAGAGTGGCGAGGCGGTGGAGCGGC 188
QY 213 GCAGACCCCTGCAGTGGCGGTGCGGAATCCGGCTCCAAAGGCGCAGAACTGTGTAGACT 272
Db 189 GGGGGGGTGGCTGCGCGGGCGGCTGGCTGGCCAGCCCGCGGCAAGACGGCT 248
QY 273 -----CACCATGTGTGCTGCCCAACGACATGCTGCACGATCCGATCCTGCGGGC 320
Db 249 GCAACGGCTGACGAACAACGTCGACGCCCTGGCGGAGGACTGCGCGGGGAGCTGCTGC 308
QY 321 GGGCAGCGGCTGCACTGGACAGGGGGGGGTGGACTTCTCGACGCCAACCCGTTCCG 380
Db 309 CGCGCGGGCTTCCGGCTGAATGCAATCGCGGGCGCGCGCTGAATCTGCTCCG 368
QY 381 CGAGTGGGCGCATCTGGCGCCAGCTCTGCGCGGAGGACAGAAACGGGGGAGTGTCT 440
Db 369 GGAATGGCAGACTGCTGGAGCAGTGTCTCGGAGGCGCGCGAGGAGGGCTGTGC 428
QY 441 TCGCAGCTGAACCCGAAAGAACCGGCGAATTTCATGCTGGCTGTTCACCGGGCTCCA 500
Db 429 CGCGCGCTCGTCCGCGCGACACGCTCAGCGCGGTGGTGGCGCGACACCGGTTTCA 488
QY 501 GCGGCTCTCCGGGTCACTTCGACCGCCAGGACCTGCGGCACCGGATCTCGTGTATG 560
Db 489 ACTGCTCGCGCGCGGACCCCGAGT---GCTCTCCGGCCAGTCTGCGCGCGTTCTG 545
QY 561 GAACACGCTGTCTCCAGCATCTGCGCGGCTCATGCTGACCTGATCGAAACCGGCGA 620
Db 546 GCGGTTACTGCTCGCGCGGCGGCGACGCGCGGCGCTTACCGCGGTGACCCGAGCG 605
QY 621 GGAGCGGATCGGAAGTTCGCGCGCGCGCGCGGCGGCGGCGCGGCTG 667
Db 606 GACGTCCCGCGCGGCGAGACGCGGACCCCGCGGACCAACCGCG 652

RESULT 7

US-156-761-2257
; Sequence 2257, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2257
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(654)
US-156-761-2257

Query Match 13.4%; Score 93.4; DB 15; Length 654;
Best Local Similarity 49.9%; Pred. No. 9.2e-15;
Matches 281; Conservative 0; Mismatches 261; Indels 21; Gaps 1;
QY 34 CAGGAACGGCGCTCGCAGCGGCGAGGCGATCGTGGGGGAGCGCGCTCGGTCTTCGAC 93

Db 85 CAGCAGCGTGCAGATCCGACCCCGGGCACGATCTCTCAACGCGCGCCCGCGGGGTTCGCC 144
QY 94 GAGTACGGGTTCAGAGCCGCCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTCAACAAG 153
Db 145 ACGGACGGCTTCCCCCAGGTACCATCAAGGACATCGCCGACGGCGCGGAGATGACCAAG 204
QY 154 GGGCGATGTACTTTCACATTCGCTTCCAAAGGAAGAGTGGCGCGCGGCTGTGCGCGGAG 213
Db 205 GGAGCGGTCTACTTTCACATTCCTCCCAACAGAGGCGCTCGCGGTAGCGGTGCTGGAGGAG 264
QY 214 CAGACCCCTGCAGTGGCGGTGCGGAATCCGGCTCCAAAGGCGCA----- 257
Db 265 TTCTACCGCGCATGAGAGGCGCTCAACGGCGCTCTCGAGCA CGGGGACCCCAATCC 324
QY 258 -----GGAACCTGTGTAGACTTACCATGTGTGTCGCCACGCGATGTGACGATCCGATC 312
Db 325 CGACGACCGTCTGTCAGCTGATAGGCGCTTCCCGCGCGCTTCCACGAAGACGTCTTC 384
QY 313 CTGCGGGCGGGCAGCGGGCTCGCACTGGACACAGGGGGCGGTGGAATCTTCCGACGCCAAC 372
Db 385 ATCCACGCGCGGGCGCGCTGCAGATCGAAACGCGCGGTACATCAAGGCCGAGCTGCCGGTG 444
QY 373 CGGTTTCGGGAGTGGGGCGACATCTGCGCCAGCTCTGCGGAGGCACAGGAACGGGG 432
Db 445 CGGTAGTGGGAGCGTGAAGTGTCTACCGAGCTCTCTGATCATGTCGCGCACCGCGCGC 504
QY 433 GAGGTGCTTTCGACGTGAACCCGAAAGACCGCGACTTTCATCGTCTGGCTGCTTCACC 492
Db 505 AACCTCCCAAGACGACGAAGCCCGAGGCGCTTGGCGCGCTCGGCTCCGCGGTTC 564
QY 493 GGGCTTCAGCGGCTTCCCGGTCTACCTCGACCGCGCAGGACCTCGGCGCACCGGATCTCG 552
Db 565 GCGCGCCACACATCTCTGGGTGTGAACGACCGCGCGGAGACATCTGGAGCGGGTGAG 624
QY 553 GTGATGTGGAACACGCTGCTGCC 575
Db 625 GAGATCATCGACGCTTCTGTGCC 647

RESULT 8

US-10-168-663-18
; Sequence 18, Application US/10168663
; Publication No. US2004008692A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909US0
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 9521
; TYPE: DNA
; ORGANISM: Streptomyces coelicolor
US-10-168-663-18

Query Match 12.3%; Score 85.8; DB 17; Length 9521;
Best Local Similarity 50.1%; Pred. No. 7.3e-13;
Matches 281; Conservative 0; Mismatches 262; Indels 18; Gaps 2;
QY 30 GCGACAGGAACGGGCGTCCGACCGCGCAGCGGATCGTGGCGGAGCGCGCTCGGTCTT 89
Db 4872 GCAGCAGGAGCGGTGATCAAGACCGCGGCCAGATCTTGGAGGCGCGCTCGGAGATCTT 4931

90 CGACGAGTACGGGTTCGAGGCGCGCCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTAC 149
 4932 CGCTCGCGCGGTACCGAGGGGCTTCGTCAAGGACGTTCGCGAGCGTGTGCGGATGAC 4991
 150 CAAGGGCGGATGATCTTCATTCGCTTCAAGGAGAGTGGCGCGGCGGTCTGCGC 209
 4992 CAAGGGCGGATGATCTTCATTCGCTTCAAGGAGAGTGGCGCGGCGGTCTGCGC 5051
 210 CGACGAGACCTGACGCTGGCGG-----TGCCGGAATCCGGCTCCCAAGGC----- 254
 5052 GGAGCACTACGCGCGTGGCGCGAGCGAGTGAAGAGATCCCAATCCAGGGCTTCACACC 5111
 255 ---CGAGGAATCTGGTAGACTCACCATGCTGGTCCGCCACGCGATGCTGACGATCCGAT 311
 5112 GCTGAGACGCTCGAGGAGATGCTCCATCGCGCGGCGAGGCTTCCCGGACGACCCCGT 5171
 312 CTTGGGGGGGGGCGACCGGCTCGCATGAGACAGGGGGGGTGGAGATCTTCGACGCCAA 371
 5172 GATGAGGCGGCTGGCGGCTCGAGAGTGAAGCGCGCTTTCATCGACGCGGAGCTGCCCT 5231
 372 CCGGTTGCGGAGTGGGCGGACATCTGCGCCCGAGCTTCTGGCGGAGGACACAGAAACGGG 431
 5232 GCGCTACGAGGAGTGAACCCACTGCTGAGGTGCGGTTGAGAGAGCCCGTGGGCGCG 5291
 432 GGAGGTGCTTCCGCGACGTGAACCCGAAAGAGCCGCGACTTCATGTCGGTGTCTAC 491
 5292 CCAATTGCGGGCGGCGTGCATCCCGACAGCTGCGCGTTCCTGTTGCGCGCTTCTT 5351
 492 CGGGTTCAGGGCGTCTCCGGGTACCTCCGACCGCAGAGACTCGGCGACCGGATCTC 551
 5352 CGGCATGACGACGCTCTCCGACAAATCTGCACACGAGCGGACATCATGAGCGGTGGCA 5411
 552 GGTGATGTGAACACGCTGCT 572
 5412 GGAGCTGCGGAGCTGATGTT 5432

RESULT 9
 US-10-168-663-19/c
 ; Sequence 19, Application US/10168663
 ; Publication No. US2004008692A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Plant Biotechnology Limited
 ; APPLICANT: Chater, Keith F
 ; APPLICANT: Bruton, Celia J
 ; APPLICANT: O'Rourke, Sean J
 ; APPLICANT: Wietzorrek, Andreas W
 ; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
 ; FILE REFERENCE: 0380-P02909USO
 ; CURRENT APPLICATION NUMBER: US/10/168,663
 ; PRIOR FILING DATE: 2002-10-25
 ; PRIOR APPLICATION NUMBER: PCT/GB00/04972
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: GB 9930477.6
 ; PRIOR FILING DATE: 1999-12-23
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 19
 ; LENGTH: 9521
 ; TYPE: DNA
 ; ORGANISM: Streptomyces coelicolor
 US-10-168-663-19

Query Match 12.3%; Score 85.8; DB 17; Length 9521;
 Best Local Similarity 50.1%; Pred. No. 7.3e-13;
 Matches 281; Conservative 0; Mismatches 262; Indels 18; Gaps 2;
 30 GCGACAGGAACGGGCGCTCCGCGAGGCGGATCGTGGCGGCGAGCCGCTCGGTCTT 89
 4650 GCAGCAGGAGCGGTTCGATCAAGACCCGCGGCCAGATCCTGGAGCGCGCTCGGAGATCTT 4591
 90 CGACGAGTACGGGTTCGAGGCGCGCCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTAC 149

Db 4590 CGCTCGCGGCGCTTACCGAGGGGCGCTCGGTCAAGGAGCTTCCCGAGCGGTGTCGGCATGAC 4531
 QY 150 CAAGGGCGGATGATCTTCATTCGCTTCAAGGAGAGTGGCGCGGCGGTGCTGCGC 209
 Db 4530 CAAGGGCGGATGATCTTCATTCGCTTCCCGAGCAAGGAATCACTGGCCATCGCGGTGGTGA 4471
 QY 210 CGACGAGACCTGACGCTGGCGG-----TGCCGGAATCCGGCTCCCAAGGC----- 254
 Db 4470 GGAGCACTACGCGCGCTGGCGCGAGCGATGGAAGAGATCCCGATCCAGGGCTTCACACC 4411
 QY 255 ---CGAGGAATCTGGTAGACTCACCATGCTGGTCCGCCACGCGATGCTGACGATCCGAT 311
 Db 4410 GCTGAGACGCTCGAGGAGATGCTCCATCGCGCGGCGAGGCTTCCCGAGCGACCCCGT 4351
 QY 312 CTTGGGGGGGGGCGACCGGCTCGCATGAGACAGGGGGCGGTGGAGATCTTCGACGCCAA 371
 Db 4350 GATGAGGCGGCTGGCGGCTCGAGAGTGAAGCGCGCTTTCATCGACGCGGAGCTGCCCT 4291
 QY 372 CCGGTTGCGGAGTGGGCGGACATCTGCGCCCGAGCTTCTGGCGGAGGACACAGAAACGGG 431
 Db 4290 GCGCTACGAGTGAACCCACTGCTGAGGTGCGGTTGAGAGCGCGCTGAGGCGCG 4231
 QY 432 GGAGGTGCTTCCGCGACGTGAACCCGAAAGAGCCGCGACTTCATGCTGGTGTCTTAC 491
 Db 4230 CCAATTGCGGGCGGCGTGCATCCCGACAGCTGCGCGTTCCTGTTGCGCGCTTCTT 4171
 QY 492 CGGGTTCAGGGCGTCTCCGGGTACCTCCGACCGCAGAGACTCGGCGACCGGATCTC 551
 Db 4170 CGGCATGACGACGCTCTCCGACAAATCTGCACACGAGCGGACATCATGAGCGGTGGCA 4111
 QY 552 GGTGATGTGAACACGCTGCT 572
 Db 4110 GGAGCTGCGGAGCTGATGTT 4090

RESULT 10
 US-10-156-761-3690
 ; Sequence 3690, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 3690
 ; LENGTH: 861
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(861)
 US-10-156-761-3690

Query Match 11.2%; Score 78.6; DB 15; Length 861;
 Best Local Similarity 58.2%; Pred. No. 6.3e-11;
 Matches 138; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 28 GTGCGCAGGAACGGGCGCTTCGCGAGCGGCGGCGGATCGTGGCGGCGAGCCGCTCGGTC 87
 Db 13 GTCAGACAGGAACGGCGGATCCGCGACGCGCAACGCTTGTATCGAGTCCGCGCGCGGCTC 72

QY 88 TTCGACGAGTACGGGTTCGAGGCGCCGACAGTGGCAGAGATCTCTCGGGGCTCGGTTC 147
Db 73 TTCGCCCGGACGGGTACGAAAGTGGTCTCGCTTCGACGATCAGCCGACGAGCGGGGTC 132
QY 148 ACCAAGGGCGGATGATCTTCCATCTCGCTTCCAAAGGAAGAGCTGGCCCGCGCGGTGCTG 207
Db 133 AGCAACGGTGGCTGACATCTCCATTTCCTCCAGCAAGGCCCTTGGCCGAGGCGGTTCGGG 192
QY 208 GCGGACGAGACCTGTGACGCTGGCGGTGCGGGAATCCGGTCCAAAGCGGAGGAATG 264
Db 193 GTGGCGGCGGAGCGATTCGGGGCGATCATCGCGGCGAGGCGCTGCGCGGCTCAG 249

RESULT 11

US-09-825-288A-1
; Sequence 1, Application US/09825288A
; Publication No. US20020192822A1
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317USC1
; CURRENT APPLICATION NUMBER: US/09/825,288A
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 09/259,821
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-1

Query Match 11.2%; Score 78.2; DB 9; Length 4257;
Best Local Similarity 45.0%; Pred. No. 7.2e-11;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY 43 GCGGTCCGACGCGGCGAGGCGATCGTGGGGCAGCGCGCTCGGTCTTCGACGAGTACGGG 102
Db 1867 GCGGTCCGCGCGCGCGACCGGGCGAGCGCGGTCTCCCGCGGTACGGCGCGCGGG 1926
QY 103 TTGAGGCGCCACATGGCAGAGATCTTCTCGCGGCGCTCGGTTCACCAAGGCGCGGATG 162
Db 1927 ATCTCGCGCGCTGGGGCGGCTGTCTCGCGCGCGCTCCCGCGGGGCGCGAC 1986
QY 163 TACTTCCACTTCGCTTCCAAAGGAAGAGCTGGCCCGCGGCTGTGGCCGAGCAGACCTG 222
Db 1987 CCGGACGCGCGCGCGCACGCGCGACGCGAGCGCGGGCGCGCGCCCGCCAGCGCGG 2046
QY 223 CAGTGGCGGTGCGGAAATCCGGCTCCAGGGCGCAGGAATGCTAGACTCACCATGCTG 282
Db 2047 CGGTGGCGGTGAGTGCCTGCGCGCTGCGCGGGATCTTGAGGGGTGCGCGGAGGGC 2106
QY 283 GTGCGCCA CGGATGTGTCACGATCCGATCTCGCGGGCGGGACGCGGGTTCGACATGAG 342
Db 2107 TTGACGCGGACCTGGCGGCGCTCCCGGGCTGGCGGGCGCGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGGCGGTGACATCTTCGAGCGCAACCCGTTTCGGCGAGTGGGGCGACATCTCGGC 402
Db 2167 CGGCGGAGGAGACCGCGGGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTGCGGAGGACACAGGAACCGGGGAGGTGCTTCCGACGCTGAACCGGAAAG 462
Db 2227 CTGCGCGGTGGTGGCGGAGTGGCGGTTCGTGCGGACGCGGTGTGCTATGCGCTG 2286
QY 463 ACCGCGACTTATCTGTGCGCTCTTACCGGGCTCCAGCGGTCTCCCGGGTCACTCC 522
Db 2287 CGGCGGAGACTGCGCGCTGGCGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 2346
QY 523 GACCGCGAGGACCTCGGCGACCGGATCTCGGTGATGTGGNAACCAAGCTGTGCGCCAG 582

Db 2347 AGCTTGTGCGCGGGCGCTTGGGCCCTCGCGGACCCGCGCTCCCGGCTTCGCGAGTCC 2406
QY 583 GTGCCCGGTTCATGCTGACCTGGAATCGAAACCGGCGAGAGCGGATCGGGAAGGTGCGG 642
Db 2407 GCGCGCGCGCGCGCGCGGACCTGTGTTTGAACAACAGAGCTTGGCGCCCTGCTGGCG 2466
QY 643 GCGCGCGCGAGGCGCGCGGAGGCTGCGAGGCTTCCGAGGCGCGCTTCGAG 693
Db 2467 GCGCGCGCGAGGCGACCGGACCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 2517

RESULT 12

US-10-156-761-2886
; Sequence 2886, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2886
; LENGTH: 18438
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(18438)
US-10-156-761-2886

Query Match 10.0%; Score 69.6; DB 15; Length 18438;
Best Local Similarity 47.3%; Pred. No. 1.1e-08;
Matches 281; Conservative 0; Mismatches 304; Indels 9; Gaps 2;

QY 99 CGGGTTCGAGGCGCGCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTTCACCAAGGGCGC 158
Db 1467 CCGGTTCGCGAGCGGACACCGCGCGGACCTCTGTGCGGCGACCGGTCTCCCGGGCGC 1526
QY 159 GATGTAATTTCCATCTTCGCTTCCAAAGGAAGAGTGGCCCGCGGCTGTGGCCGAGCAGAC 218
Db 1527 GCTCGGTTTGTCTTTCACCGGCGGGCAGCAGCGCGCGCATGACCGCGGAGCTGA 1586
QY 219 CTGCGAGTGGCGGTGCCGATCCGCTTCCAAAGGCGGAGGAGTGGTAGACTCACCAT 278
Db 1587 CCGGAGCTCCGCGGAGTACG---CCGCCCTCTTGAAGGTTGCGCGAGCTTGACCC 1643
QY 279 GCTGTGCGCCCGGCGATCTGCACGATCCGATCTCGCGCGGCGGACGCGGCTCGCACT 338
Db 1644 GTGGTTCGACGCGCGCTGCGGACGCTGTGTTCGCGCGGAGGTTCCGACGAGGCGGC 1703
QY 339 GGAACGAGGGGCGGTGGACTTTCGACGCGCAACCGTTTCGCGAGTGGGGCGACATCTG 398
Db 1704 GCTCTGGACCGGTACCGAGTTTCAACCGCGCGCTTTCGCGCTCGAGGTCCGCTCTT 1763
QY 399 CGGCCAGCTCTTGGGGGAGGACAGGAAAGGGGAGGTGCTTCCGACGAGTGAACCGGAA 458
Db 1764 CCGGTACGCGGAGCACTGCGGCGCTGACACCGCGGTTCCTGCTCGGTCACTCGGTG 1823
QY 459 AAGACCGGCGACTTTCATCTGCTGGCTGTTCACCGGGCTCCAGGCGGTCTCCCGGGTCA 518
Db 1824 ACTGGGCGCGCGACGCTGCGCGAGTCTCTCTCCCTGGGGACGCTGACGCGTCTGCTGC 1883

Qy	519	CTCCGACCGCCAGGACCTCGGCCACCG-----GATCTCGGTGATGTGGAAACACCGTGT	572
Db	1884	CGCCCGCGGCGGCTGATGCAGGACCGCGCGGCGACCGGCGCATGTGTCTCGCTCCAGGC	1943
Qy	573	GCCGACGATCGTCCCGCGTCCATGCTGACTGGATCGAAACCGCGGAGAGCGGATCGG	632
Db	1944	CACCGAGGACGAGCTGGCAACGTTCTCTCCGAGTCGGTACCCCTGGCGCGCTGACCG	2003
Qy	633	GAAGGTCCGCGCGCGCCGAGGCGCGGAGGCTCGCGAGGCCCTCGAGAGCCCG	686
Db	2004	CCCCGCTCCACGGTGATCCGCGGTGACGAGGACGCCGTACTGGCGCTCGCCGC	2057

```

RESULT 13
US-10-156-761-15102
; Sequence 15102, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 15102
; LENGTH: 125746
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
US-10-156-761-15102

```

Query Match	10.0%	Score 69.6	DB 15	Length 125746
Best Local Similarity	47.3%	Pred. No. 9.9e-09		
Matches 281	Conservative 0	Mismatches 304	Indels 9	Gaps 2
Qy	99	CGGTTTCGAGGCGGCACACAGTGCACAGATCTCTCGGGGCTCGGTCAACCAAGGGCGC	158	
Db	26777	CCGCTCGCCGACGGCACACCCGGCGCGAACCTCTGTCGGGACGGTCTCCCGGGCGC	26836	
Qy	159	GATGTACTTCCACTTCGCTTTCAGGAAGAGCTGGCCCGCGCTGCTGGCGGACAGAC	218	
Db	26837	GCTCGCTTCGTCTTTCACCGGGCAGGCGACGACGCCCGGCATGACCGCCGAGCTGTA	26896	
Qy	219	CTGCGAGTGGCGGTGCCGGAATCCGGCTTCAAGCGCAGGAACTGTTAGACTCTACCAT	278	
Db	26897	CCGAGCTCCGCGAGTACG---CCGCGGCTTGGACGAGGTGTGCGGAGGCTGGACCC	26953	
Qy	279	GCTGGTCGCCACGGCATGTGACGATCCGATCTCTCGGGGGGGGACGCGGCTCGCACT	338	
Db	26954	GTGACTGGACGGCGCTGCGGACGTGCTGTTTCGCGCGGAAGTTCCGACGAGGCGGC	27013	
Qy	339	GGACGAGGGGGGTGACACTCTTCGACGCCAACCCGTTTCGGCGAGTGGGGCGACATCTG	398	
Db	27014	GCTCTTGGAACGTACCGAGTTTCAACGACCGCCCTGTTTCGCGGTGAGGTGCGCCTCTT	27073	
Qy	399	CGCCACGCTCTGCGGAGGACAGGAACGGGGGAGGTGCTTCCGACGTGAACCCGAA	458	
Db	27074	CCGTTACGCCGAGCATCGCGCTTGACACGCGGTTCCTGCTCGGTCACTCGGTGGTGA	27133	
Qy	459	AAAGACGGGACTTTCATCGTGGCTGCTTTCACCGGGCTCCAGGCGGTCTCCCGGGTCAAC	518	
Db	27134	ACTGGCGGCGCGACCGTTCGCGGAGTCTCTCCCTGGGGGACGCTGACGCTGCTGCGC	27193	

Qy	519	CTCCGACGGCCAGACCTCGGCCACCG-----GATCTCGGTGATGTGGACACCATGTGCT	572
Db	27194	CGCCCGCGCGCGGTGATGTCAGGACCGCGCCGCGACCGCGCGATGGTGTTCGGTCCAGGC	27253
Qy	573	GCCGAGCATGTCGCGGCGTCCATCTGACCTGGATCGAAACGCGGCGAGGAGCGGATCGG	632
Db	27254	CACGAGGACGAGCTGGCACCCTTCGTTCTCCAGTTCGGTACCTCGCGCGCTGAACGG	27313
Qy	633	GAAGGTGCGCGCGCGCGCGGAGCGCCGAGGCTCGGAGGCTCCGAGGCGCGC	686
Db	27314	CCCCGCTCCACGGTGATCGCGGTGACGAGGACGCGTACTGGCCGCTGCGCGC	27367

RESULT 14
 US-10-437-963-77080
 ; Sequence 77080, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Boukharov, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 77080
 ; LENGTH: 3173
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_77012C.1
 ; US-10-437-963-77080

RESULT 15

US-10-282-122A-25841
; Sequence 25841, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25841
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mycobacterium avium
US-10-282-122A-25841

Query Match 9.8%; Score 68.2; DB 17; Length 1224;
Best Local Similarity 46.4%; Pred. No. 3.1e-08;
Matches 306; Conservative 0; Mismatches 338; Indels 15; Gaps 2;
QY 36 GGNACGGGCGTCCGACGGGAGCGGATCGTGGCGGAGCGCCCTCGGTCTTCACGA 95
DB 522 GCAGCGGGTGCTGTCCCGGTGGCCGATTTCGGCCCTGCCCGCGCGAGGTGTGGCGT 581
QY 96 GTACGGGTTTCGAGCGCGCCACAGTGGCAGAGATCTCTCGCGGCGCTCGGTACCAAGG 155
DB 582 CCGGTGCGGAGGCGCATACCTCGGTGACCGCGCGGCGGTGCGCGCGCGGTCTG 641
QY 156 CGGATGTATCTTCCTTCCTTCCAAAGGAAGCTGGCCCGCGGCGTGTGGCCGAGCA 215
DB 642 GCACCTTCGATTTCACGACGAGCCCGGGTGTGTGTTTCGGCGGCTCGCAGGGCGC 701
QY 216 GACCTTGACGTGGCGGTGCGGAATCCGGCTCCAAAGGCGCAGGAACCTGTAGACTAC 275
DB 702 GGCTTCGCTGAACCGGGCGGTGTTCGGCGCGCGCGAGCTGGCGCGCGCGCGGTGTC 761
QY 276 CATGCTGGTCGCCCGCGGATGTGCGATCCGATCCGATCCGCGCGCGCGCGGTGTC 335
DB 762 CGTGTGTGACGGCGCAGCGGCCCAAAACACGCTCGACCTCGCGCAACCGCAGCCCGCGCA 821

QY 336 ACTGGACCAAGGGCGGTGGACTTCTCCGAC-----GCCAACCCGTTCCGCGA 383
DB 822 CCGCGCCTACGTGGCGGTGCCCTACCTGGACCGAATGGACCTGCTACGCGCGCGCGA 881
QY 384 GTGGGGGACATCTGCGCCAGCTCCTGGCGGAGGACAGGAACGGGGGAGGTGCTTCC 443
DB 882 CTTGGTGATCTGCCGGTCCGGGGCGATGACGGTCCGCGAGGTGTCCGCGGTGGGGTTGCC 941
QY 444 GCACGTGAACCCGAAAAAGACCGGCGACTTTCATCGTGGTCTTCAACCGGGCTCCAGGC 503
DB 942 GGCCATCTAGTGGCGCTGCCGATCGGCAACGGCGAACAGCGGTCAACCGGTGCGGT 1001
QY 504 GGTCTCCGGGTCACTCTCCGACCGCCAGGACTTCGCCCAACCGGATCTCGGTGATGTGAA 563
DB 1002 GGTCAACGGCGGCGGCGCATGGTGGTTCGCCGACGCGGCTTACGCGCGAGTTGGTTCG 1061
QY 564 CCACGTGCTGCCAGCATCGTCCCG--CGTCCATGCTGACCTGGATCGAAGCGCGA 620
DB 1062 CCGGGAGGTGGCGCGCTGGTGGCGACCGCGCGCGCTGGCGCGATGACGACGCGCGC 1121
QY 621 GGAGCGGATCGGGAAGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 679
DB 1122 CGCGCGGTTCGGGCACTCCCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1180

Search completed: March 6, 2005, 22:27:00
Job time : 657 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 6, 2005, 19:46:55 ; Search time 3469 Seconds
(without alignments)
7669.912 Million cell updates/sec

Title: US-10-049-710A-1
Perfect score: 699
Sequence: 1 atggcagtcgacacgaacg.....aggccgctccgacgagtag 699

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69.2	9.9	732	8	BZ564850 pacs2-164
2	68.4	9.8	925	9	CNS0091P
3	67.2	9.6	967	8	BZ567330
4	66	9.4	1418	3	AK019388 Mus muscu
5	65	9.3	783	9	CG450596 OGVGUS2TV
6	65	9.3	822	9	CG444666 OGVG35TH
7	64.6	9.2	925	9	CNS0091P
8	63.8	9.1	1016	7	CF583098
9	63.6	9.1	846	7	CN146323 AGENCOURT
10	60.8	8.7	584	7	CK626160 WOUND1.39
11	60	8.6	657	7	CN011285 WHE3882.C
12	60	8.6	935	7	CNS006XK
13	59.8	8.6	1084	7	CK213336 FGAS02524
14	59.2	8.5	483	8	AZ212924 Sheared D
15	59	8.4	640	2	BE363278 WS1_61.D0
16	58.6	8.4	2332	9	AG363333 Mus muscu
17	58.4	8.4	876	9	CC643671 OGNAA29TH
18	58.2	8.3	524	9	CG256542 OGLC26TH
19	58.2	8.3	744	9	CG322461 OGA2A31TH
20	58.2	8.3	819	8	BZ649366 OGCBY29TC
21	58.2	8.3	900	9	CC682633 OGOAK35TV
22	58	8.3	899	9	CC620156 OGUBL90TV
23	58	8.3	932	9	CNS0072Q
24	57.6	8.2	752	7	CK432635 UI-D-GC1-

ALIGNMENTS

RESULT 1
BZ564850
LOCUS BZ564850 732 bp DNA linear GSS 17-DEC-2002
DEFINITION pacs2-164_4953.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
ACCESSION BZ564850
VERSION BZ564850.1 GI:27191489
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 732)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sime, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1..732
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_4953"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match 9.9%; Score 69.2; DB 8; Length 732;
Best Local Similarity 48.0%; Pred. No. 0.00034;
Matches 197; Conservative 0; Mismatches 213; Indels 0; Gaps 0;
QY 212 AGCAGACCTGCACGTGGCGGTCCCGGAATCCGGTCCAAAGCGCGCAAGCACTGGTAGACC 271
|||
Db 141 AGCGCGCATCGAGTGTGGCGCGGTATGGCGACGAGGTGCTGGAGCGCTTCTGTCG 200
|||
QY 272 TCACCATGCTGTGCTGCCACGCGCATCTGCACGATCCGATCTCTGCGGGCGGCACCGGGC 331


```

Db      201  CCGGCGCGAGGTACCGTCGGGGTCTCGACGACGACGCGTCCGGTGGCGAGATTC 260
QY      332  TCGCACTGGACAGGGGGGGTGGACTTCTTCGACGCCAACCCGTTCCGGGAGTGGGGCG 391
Db      261  TCTCGGGCGCCAGGAGGTCTTCGACTACGAGCAAAAGTACCAGCGCGCGGTGCGCG 320
QY      392  ACATCTGCCGCCAGCTCTTCGGGGAGGACAGAGAGCGGGGAGGTCTTCGACACGTA 451
Db      321  AGGTGTTCCTCCCGACCTTCAGCGGGCGGATCGCCCGGAGGCCAGCGTCTGGCGCTGA 380
QY      452  ACCCGAAAAGACCGCGCACTTCATCGTCGGTGTCTTACCGGGGTCTCAGGCGGTCTCCC 511
Db      381  AGGTACACCGGGCGGTGAGCTGAGCGGTACAGCGCACCGACTTCGACTCGACGAAC 440
QY      512  GGGTCACTCCGACCGCCAGGACTCTCGCCACCGGATCTCGGTGATGTGGAACACGTCGC 571
Db      441  AGGGCGGGTCTGTGCTGCTGGAGGTCAATACCTCTCGCGGCGATGACCGCACCGCCTGC 500
QY      572  TGCCAGCATCGTCGGCGGTCATGCTGACTGATCGAATCGAAACCGCGCGAG 621
Db      501  TGCCGACCGCGGCGCGCGCGGATCGAGTGTCCGAACTCTGCGAG 550

```

```

RESULT 2
CNS0091P/c
LOCUS      925 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL053013.1 GI:4934461
VERSION
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="End : TET3"

```

```

FEATURES
source
Query Match      9.8%; Score 68.4; DB 9; Length 925;
Best Local Similarity 14.1%; Pred. NO. 0.0005;
Matches 55; Conservative 183; Mismatches 152; Indels 0; Gaps 0;
ORIGIN

```

```

QY      168  CCACTTCGTCTTCAAGAGAGAGCTGCCCGCGCGTCTGCCGAGCAGACCTGACGCT 227
Db      919  CSCCSBSBSCSSSMSTSSNSBSCSSBSSTSSMSBSBSBSBSBSBSBSBSBSBSBS 860
QY      228  GGCGGTCCCGGATCCGGCTCCAAAGCGCAGGAACTGGTAGACTCACCATGCTGTGTCGC 287
Db      859  ACVKCNASSCGCGCGCMABCCMSCSSSSSCCGASARGVKVRASGGAGKRGSGSGASAS 800
QY      288  CCAGCGCATGTCGACCATCCGATCCTGTCGGGGGGGCGACGCGCTCGCATGGAACACGG 347
Db      799  HSSSACBSSSSSSCASCSWSASSSSSASSSRSRSGGAGGSGASSRSSSSSSASAGSV 740
QY      348  GGCGGTGAGATTCTCCGACGCCAACCCGTTCCGGGAGTGGGGCGACATCTCGGCCAGCT 407
Db      739  VSSASSSSSSSSSSSVCSVSSVSSMSCSSBSBSBSBSBSBSBSBSBSBSBSBSBS 680
QY      408  CCTGCGGAGGACAGAGACGGGGGAGGTCTTCCGACGCTGAACCCGAAAGAACCGG 467
Db      679  CSTSASMSAARSSSSSSSSSSSSMSASSSASSSSSSSSSSSSSSSSSSSSSSSSGG 620
QY      468  CGACTTCATCGTCGGTCTTACCGGGTCCAGCGGTCTCCGGGTACCTCCGACCG 527
Db      619  GSGSVSASSGMSVSSVSSGGRSGSGGGVGGSGSSSSSGSGSGSGSVCSSSSGMCR 560
QY      528  CCAGGACCTCGGCCACCGGATCTCGGTGAT 557
Db      559  CSCSAAAAAASCAVAAACGMMCGKSKSGCT 530

```

```

RESULT 3
BZ567330
LOCUS      967 bp      DNA      linear      GSS 17-DEC-2002
DEFINITION pac82-164_6885.Y2 pac82-164 Pseudomonas aeruginosa genomic clone
            pac82-164_6885, genomic survey sequence.
ACCESSION  BZ567330.1 GI:27198447
VERSION
KEYWORDS
SOURCE      Pseudomonas aeruginosa
ORGANISM    Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 967)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence Variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..967
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac82-164_6885"
/clone_lib="pac82-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."

```

```

FEATURES
source
Query Match      9.6%; Score 67.2; DB 8; Length 967;
Best Local Similarity 46.9%; Pred. NO. 0.00088;
Matches 210; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
ORIGIN
212  AGCAGACCTTGCACGTGGCGGTGCGCGGATCCGGATCCGAGGCGGAGGAACTGGTAGACC 271

```



```
Query Match 9.4%; Score 66; DB 3; Length 1418;
Best Local Similarity 47.0%; Pred. No. 0.0015;
Matches 259; Conservative 0; Mismatches 275; Indels 8; Gaps 2;

QY 143 CGGTACCAAGGGCGGATGTTACCTTCACCTTCCTTCCAAAGGAAGAGTGGCCCGGGCG 202
D 143 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 363 CGGTGACGCGACTCGTGGTGGTGAACCCCGACGAGCGCTCCCGCTGCTTCGGGT 422
QY 203 TGTGCGCGAGACAGCCCTGACGTGGCGGTGCGGAATCCGCTCAAGGCGCAGAAC 262
D 203 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 423 TGTGACCTACTCGAACGAGGAGCGGATGCCCGCATGGCCGCTGCGCCGCGCG 482
QY 263 TGTGACCTACTCGAACGAGGAGCGGATGCCCGCATGGCCGCTGCGCCGCGCG 322
D 263 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 483 TGTGACGCAACAGGTGGAGCTGAAGCGCGCTGTGCGGGAGGATTCGGCGCGCGCG 542
QY 323 GCAACGGGCTCGACATGGACAGGGGCGGTGACATTCCTCCGACGCCAACCCGCTTCGGCG 382
D 323 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 543 GGGCGCACGCAAGGTGAAGAAGCTGTGCTGGCGCGCTCAAGGGCGAC---GTGGCG 599
QY 383 AGTGGGCGGACATCTGGCGCCAGCTCTGCGGAGGACAGGAAAGGGGGAGGTCTTC 442
D 383 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 600 AGGGACACCTGATCGAGACTTCTCGAGTTCGCGCGGTGCGGAGAGGCGGAGATCAT 659
QY 443 CGCACGTGAACCCGAAAAAGACCGGCGACTTTCATCTGCGCTGCTTTCACCGGCTCCAG 502
D 443 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 660 CCGACAAAGCAGTCGGGCAAGAGCGGCTTCGGCTTCGTCTACTTCCAGAGCCAGAG 719
QY 503 CGGTCTCCGGGTACCTCGGACCGGAGGACCTCGGCAACCGGATCTCGGTGATGTGA 562
D 503 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 720 CGGCGCAACAGGCGCGGTGAGTTCAGTTCCACCGGATCCAGGGCCACCGCGTGGAG 779
QY 563 ACCACGTGTCGCCAGCAGTCGTGCGCGCTGCATGCTGACCTGATCGAAGACCGGCGAG 622
D 563 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 780 AGAAGCGGTGCCAAG-----GAGATATCCAGCGGCGGGGGTGGCGGCGCG 834
QY 623 AGCGGATCGGGAGAGTTCGCGCGCGCGCGGAGCGCGCGGAGGCTCGGAGGCTCCGAG 682
D 623 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 835 CCGGGCGGCGCGGCGGAGGCGCGGCGCGCGCGCGGCGGCGCGCGCGCGGCGGCGAG 894
QY 683 CC 684
D 895 CC 896

RESULT 5
CG450596
LOCUS CG450596 822 bp DNA linear GSS 17-SEP-2003
DEFINITION OGVS52TV ZM 0.7_1.5 KB Zea mays genomic clone ZMMBma0479107,
genomic survey sequence.
ACCESSION CG450596
VERSION CG450596.1 GI:34835596
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 783)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGVS52TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
Source
Location/Qualifiers
1..783
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0479107"
/clone_lib="ZM 0.7_1.5 KB"
/note="Vector: pBCK; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 9.3%; Score 65; DB 9; Length 783;
Best Local Similarity 47.0%; Pred. No. 0.0025;
Matches 240; Conservative 0; Mismatches 265; Indels 6; Gaps 1;

QY 54 GCGGAGGCGGATCGTGGCGGAGCGCCCTCGGTCTTTCGACGAGTACGGGTTCGAGCGCG 113
D 54 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 208 CGACGAGCGGCGCGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267
QY 114 GACAGTGGCAGAGATCTTCTCGCGGCGCTCGGTACCAAGGGCGGATGTTACTTCACAT 173
D 114 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 268 CACCATGCAACGACGAGCTGTCGCGCGTGGCGCTCGAGAAGCGCGCATGCTCATGCTGT 327
QY 174 CGCTTCCAAAGGAAGAGCTGGCGCGGCTGCTGCGCGAGAGACCTGACGTGGCGGT 233
D 174 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 328 GCCTTCCACACGCGCTCGCATCGAGGGCTCGTTCGAGAACAACGAGCGCAACGCCGG 387
QY 234 GCGGAAATCCGGCTCCAAAGGCGCAGGAACCTGTGTAGACCTCACCATGTGTGTGTCGCG 293
D 234 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 388 CTCGCTCGAGCGCGCAACGTCACGCTCTCCACTACTCCCTG-----CTCGCTGCG 441
QY 294 CATGCTGCGACATCCGATCCTCGCGGCGGCGACCGGCTCGACCTGGAACCGAGGGCGGT 353
D 294 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 442 CATCTCTGTCGACCGGGGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
QY 354 GACATCTTCCACGCGCAACCGCTTTCGGGAGTGGGGCGACATCTCGCGCCAGCTCTCTGCG 413
D 354 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 502 GSCACAGCGCGCGCGAGCGCTTCCCTCACCGGCTCGGCTCTACTTCTCTCGCGGGGCC 561
QY 414 GAGGACAGGAAACCGGGGAGGTCTTCCGCACTGTGAACCGGAAAAAGACCGGGGACAT 473
D 414 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 562 GACGACCGGAGGCGCTCGGCTCGCACGACATGCGCGAGGACGCGGCCCATCGGCT 621
QY 474 CATCTCTGCTGCTTTCACCGGCTCCAGGCGCTCTCCCGGCTCACCTCCGACCGCAGGA 533
D 474 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 622 CACCTCTCTCGGTTCTGCGCGCGCGAGTGGCGCAAGGGCGGAGCGCGCGCGAGGA 581
QY 534 CTTGCGGCGCGGATCTCGGTGATGTGAAC 564
D 534 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 682 CCGGCTCGAGGAGGCGCTGCGAGGATAC 712

RESULT 6
CG444666/c
LOCUS CG444666 822 bp DNA linear GSS 17-SEP-2003
DEFINITION OGVS35TH ZM 0.7_1.5 KB Zea mays genomic clone ZMMBma0475E21,
genomic survey sequence.
ACCESSION CG444666
VERSION CG444666.1 GI:34826493
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 822)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGVS35TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
```

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1. 822

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBMA0475E21"

/clone_lib="ZM 0.7 1.5 kb"

/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 9.3%; Score 65; DB 9; Length 822;

Best Local Similarity 47.0%; Pred. No. 0.0025;

Matches 240; Conservative 0; Mismatches 265; Indels 6; Gaps 1;

```
QY 54 GCGGAGGCGATCGTGGGGGAGCGCCCTCGCTTCGACGAGTACGGGTTGAGCGCG 113
DB |||||
QY 525 GCAGAGCGCGCGTGGGGTCCGCGTCTCGCTCCCTACGTCATCGCGCCCTACGC 466
DB |||||
QY 114 CACAGTGGCAGAGATCCTCTCGGGGCGCTCGGTCAACAGGGCGCATGACTTCCACTT 173
DB |||||
QY 465 CACCATGCACGACGAGCTCTGCGCGTGGCGTTCGAGAGCGCGCATGCTCATGCTGT 406
DB |||||
QY 174 CGCTTCCAGAGAGAGCTGGCGCGCGCTGTGCGCGAGCAGACCTCGACGTGGCGGT 233
DB |||||
QY 405 GCCCTTCCACGCGCGCTCGCATCGACGCTCGGTCCGTCGAGAACACGACGCCAGCGCGG 346
DB |||||
QY 234 GCGGGAATCGGCTCAAGGCGCAGGAATGTTAGACTTCACTGCTGCTGCTGCCACGG 293
DB |||||
QY 345 CTCGCTCCAGGCGCGCAACGTCAACGTGTCTCACTACTCCCGT-----CTCGTCCG 292
DB |||||
QY 294 CATGCTGCACGATCCGATCTTGGGGCGGCGACGCGGTCTCGACCTCGACGAGGGGCGGT 353
DB |||||
QY 291 CATCTCTGTCAGCCGGGCGAGCTCTCGTGGATCGTCCGCGTGGCGGGGCGCGCG 232
DB |||||
QY 354 GGACTTCTCCAGCGCAACCGGTTGCGCGAGTGGGCGACATCTGCGCCCGAGCTCTGCGC 413
DB |||||
QY 231 GGCACAGCGCGACGCGCTTCCCTCACCGCGTGGGCTTACTTCTCGGCGGGC 172
DB |||||
QY 414 GGAGGACAGGAACGGGGGAGTGTCTCGGACGTGAACCCGAAAGAACCGGCGACTT 473
DB |||||
QY 171 GGACGACCGGAGGCGCTCGCGCTCGCCACGACATGGCGAGGACGCGCCCATCGGCT 112
DB |||||
QY 474 CATGCTGCGCTGCTTCCACCGGGTCCAGCGGCTCTCCCGGTCCACCTCGACCGCGCAGGA 533
DB |||||
QY 111 CACCGTCTTCGGTCTCTGCGCGCGCGAGTGGCGCAAGGGCGGCGACCCGAGAGGA 52
DB |||||
QY 534 CTTGCGCCACCGGATCTCGGTGATGTGAAC 564
DB |||||
QY 51 CCGGCTCGACGAGGCGCGTGCAGGAGTAC 21
DB |||||
```

RESULT 7

CNS0091P

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence TET3 end of BAC #

BACR19D16 of RPI-98 library from Drosophila melanogaster. (fruit

fly), genomic survey sequence.

AL053013

VERSION

AL053013.1 GI:4934461

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 925)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Oosagawa and

Aaron Mammoser in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw ap, the same strain used for the BDGP's

P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

1..925

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR19D16"

/clone_lib="RPI-98"

/note="end : TET3"

ORIGIN

Query Match

Best Local Similarity 9.2%; Score 64.6; DB 9; Length 925;

Matches 53; Conservative 183; Mismatches 132; Indels 3; Gaps 1;

```
QY 332 TCGCTGACGACGAGGGCGGTGGACTTCTCGACGCCAACCGTTCCGCGAGTGGGCG 391
DB |||||
QY 554 TSSGGYGGKSGSGBSGCCSCSSSCSCBCCCCSCSSYCCSSSSSSKCSST 613
DB |||||
QY 392 ACATCTGCGCCAGCTCTTGGCGGAGGACAGACGCGGAGGTGCTTCCGACGTGA 451
DB |||||
QY 614 SBCSCCCSKVCGTSCSSSSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 673
QY 452 ACCCGAAAGAACCGCGACTTTCATGCTGCTTTCACCGGCTCCAG---GCGGTCT 508
DB |||||
QY 674 STSASGSGWSAGGSGTGTSTSSSSSTSTSSSSVSSGSKSTBSGSSBSSGSSSS 733
QY 509 CCGGGTACCTCCGACCCGAGGACCTCGGCACCGGATCTCGGTGATGTGAACACG 568
DB |||||
QY 734 SSTSSBBSCTSTSSSSSSSTSCCTCCCSYSSSTSSSTSSSTSSSTSSSTSSSV 793
QY 569 TCGTCCGACGATCGTGGCGGCTCCATGCTGACCTGGATCGAAACCGCGAGAGCGGA 628
DB |||||
QY 794 TSSSDSTSTCCSCCYMTCTCTYBMBCTTSTCGSSSSSGKGGVTKCGCGGCGSSST 853
QY 629 TCGGAAGGTTCGCGCGCGCGGAGCGCCGAGGCTCGGAGGCTCGAGGCGCGCT 688
DB |||||
QY 854 NGBMGTTSSACSSSSSSSSVSSSKSSASSSVSSGSSGSSVSSSSASKSSSSG 913
QY 689 CCGACGAGTAG 699
DB |||||
QY 914 SGSGSGSGSVS 924
```

RESULT 8

CF583098

LOCUS

DEFINITION

AGENCOURT 11361399 updated NIH_MGC_137 Mus musculus cDNA clone

IMAGE:643215 5', mRNA sequence.

CF583098

ACCESSION

VERSION

CF583098.1 GI:35196360

CNS0091P 925 bp DNA linear GSS 03-JUN-1999

Drosophila melanogaster genome survey sequence TET3 end of BAC #

BACR19D16 of RPI-98 library from Drosophila melanogaster. (fruit

fly), genomic survey sequence.

AL053013

VERSION

AL053013.1 GI:4934461

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1016)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gerard Gradwohl (PNAS 97 p1607-1611, 2000)
cDNA Library Preparation: Catherine Lee, Endocrine Pancreas Consortium
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://imgc.llnl.gov>
Plate: IRBD24 row: d column: 12
High quality sequence start: 25
High quality sequence stop: 532.

FEATURES

Location/Qualifiers
1..1016

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6432215"
/lab_host="DH10B"
/clone_lib="NIH_MGC_137"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Sall;
Site 2: NotI; Library consists of a pool of clones
rearrayed from the following libraries: Melton normalized
mixed mouse pancreas 1 NI-MMS1, Amplified Melton mouse
islets 1 M1S1-A, and Kaestner pgn3 wt. Clones rearrayed in
the laboratory of K. Kaestner (University of
Pennsylvania). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 9.1%; Score 63.8; DB 7; Length 1016;
Best Local Similarity 48.7%; Pred. No. 0.0044;
Matches 237; Conservative 0; Mismatches 242; Indels 8; Gaps 2;

QY 198 CGCGTGTGGCGGACAGACCTGACGTGGGGTCCGGATCCGGTCCAGCGCA 257
DB 68 CGCGTTCGTGACCTACTCGAACGTGGTGGAGCGGATGCCCGCGTCCGCGCA 127
QY 258 GGAACCTGGTAGACCTCACCATCTGTGTCGCCACGCGATGTCACGATCCGATCTGGG 317
DB 128 CGCGTGGAGCGCAACACAGTGGAGCTGAAGCGCGCGTGTCCGGGAGGATTCGGCGCG 187
QY 318 GCGCGGACACCGGCTGGCACTGGACACAGGGCGGTGGACTTCTCCGACCCAAACCGGTT 377
DB 188 GCCTCGGGCGCGCCCAAGTGAAGAGCTGTTCGTGGCGCGCTCAAGGGCGAC---GT 244
QY 378 CGCGGAGTGGCGGACATCTGCGCCAGCTCTGCGGAGGACACAGGACGGGGAGGT 437
DB 245 GCGGAGGCGGACCTGATCGACACTTCTCGCAGTTCGCGCGGGTGGAGAGCGGAGAT 304
QY 438 GCTTCGCGACGTGAACCCGAAAAGACCGCGACTTTCATCTCGGTGCTTCCACCGGGCT 497
DB 305 CATTCGCGACAGCAGTCGCGCAAGAGCGCGCTTCGGCTTCGTCTACTTCCAGAGCA 364
QY 498 CAGGGGCTTCGCGGTCCTCGACCTCGACCGCGGACCTCGGCGCACCGGATTCGGGTGAT 557
DB 365 CGACGCGCGCGACAGCGCGCGTGTGTCAGTTCACCGCGATCCAGGGCCACCGGTGA 424
QY 558 GTGGAAACACGTGCTGCCAGCATCGTCCGCGCTCCATGCTGATCGATGAACCGG 617
DB 425 GGTGAAGAGGCGGTGCCCA-----AGGAGGATATCCACGCGGCGCGGGGTGCGCGG 479

QY 618 CGAGGAGCGATCGGAGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGGAGGCCTC 677
DB 480 GCGGCG 539
QY 678 CGAGGCGC 684
DB 540 CGAGGACC 546

RESULT 9
LOCUS CN146323
DEFINITION WOUND1_39_C06_g1_A002 Wounded leaves Sorghum bicolor cDNA clone
WOUND1_39_C06_A002 5', mRNA sequence.
ACCESSION CN146323
VERSION CN146323.1 GI:45986172
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 846)

REFERENCE Cordonnier-Pratt M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,
Sun, F., Sullivan, R., Brady, J., Eastman, A., Miller, V., Gonzalez, M.,
Anfuso, C., Chhabra, D., Johnson, H., Kamran, D. and Pratt, L.H.
A Sorghum EST database: mechanically damaged and methyl
jasmonate-treated leaves

Unpublished (2003)
Other ESTs: WOUND1_39_C06_b1_A002
Contact: Cordonnier-Pratt M.-M.
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.

Seq primer: Sug5 (CTTCGCTCTAAAGCTGCG).
Location/Qualifiers
1..846

/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone="WOUND1_39_C06_A002"
/lab_host="DH10B-F1 phage-resistant E. coli"
/clone_lib="Wounded leaves"
/note="Organ: Leaf; Vector: pME18S-FL3; Site 1: XhoI;
Site 2: XhoI; The library was prepared from polyA+ RNA
harvested from 8-day-old hydroponically grown, BTx623
sorghum seedlings. For some plants, one-half of the second
leaf was crushed without damaging the midvein. For others,
methyl jasmonate was added to the growth medium to a final
concentration of 100 uM. Leaves were harvested 3 and 27 hr
after treatment and pooled. Double-stranded cDNA was
cloned unidirectionally into different DraIII sites of the
pME18S-FL3 vector (5'-prime DraIII site is CACTGTGTG,
3'-prime DraIII site is CACCAATGTG). XhoI excises the cDNA
insert."

FEATURES
Location/Qualifiers
1..846

/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone="WOUND1_39_C06_A002"
/lab_host="DH10B-F1 phage-resistant E. coli"
/clone_lib="Wounded leaves"
/note="Organ: Leaf; Vector: pME18S-FL3; Site 1: XhoI;
Site 2: XhoI; The library was prepared from polyA+ RNA
harvested from 8-day-old hydroponically grown, BTx623
sorghum seedlings. For some plants, one-half of the second
leaf was crushed without damaging the midvein. For others,
methyl jasmonate was added to the growth medium to a final
concentration of 100 uM. Leaves were harvested 3 and 27 hr
after treatment and pooled. Double-stranded cDNA was
cloned unidirectionally into different DraIII sites of the
pME18S-FL3 vector (5'-prime DraIII site is CACTGTGTG,
3'-prime DraIII site is CACCAATGTG). XhoI excises the cDNA
insert."

Query Match 9.1%; Score 63.6; DB 7; Length 846;
Best Local Similarity 47.9%; Pred. No. 0.0049;
Matches 183; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

ORIGIN
Query Match 9.1%; Score 63.6; DB 7; Length 846;
Best Local Similarity 47.9%; Pred. No. 0.0049;
Matches 183; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

ORIGIN
Query Match 9.1%; Score 63.6; DB 7; Length 846;
Best Local Similarity 47.9%; Pred. No. 0.0049;
Matches 183; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

ORIGIN
Query Match 9.1%; Score 63.6; DB 7; Length 846;
Best Local Similarity 47.9%; Pred. No. 0.0049;
Matches 183; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

ORIGIN
Query Match 9.1%; Score 63.6; DB 7; Length 846;
Best Local Similarity 47.9%; Pred. No. 0.0049;
Matches 183; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

Qy	281	TTGTCGCCACGGGATGCTGCACGATCCGATCCTCGGGCGGGCAGCGGGCTCGCACTGG	340
Db	171	TGTCGACACCGGCAACTGTGCGTGAACGGTCCCGCCGCGGCTCACTCTCAACCCAGC	230
Qy	341	ACCAGGGGCGGTGACATTCTCCGACGGCAACCCGTTGCGCGAGTGGGCGCACATCTGGC	400
Db	231	CCCCCTAGCGGAGACCTTCTTCGGCACTAAGCCGTACAACGCAACGAGCTTCAACGAGT	290
Qy	401	CCGAGCTCCTGGCGAGGACACAGGAACGGGGGGAGGTGCTTCGGACGTGAACCCGAAAA	460
Db	291	ACGTGCTGGAGGGGAGCGCGCGTGTCTGCGCGAGTCGCGACGGCTCTTTCAAGAAAGTCA	350
Qy	461	AGACGGCGACTTCATCTCGGCTGCTTACCGGGCTCCAGCGGCTCTCCCGGTCACCT	520
Db	351	TCTCGCTCGGCATCAAGCCGGTGTCTTGACGGGGGGCACCGAGAACGAGGGGCCATCA	410
Qy	521	CCGACCGCAGGACTTCGGGCACCGGATCTCGGTGATGTGGAAACCACTGCTGCCGAGCA	580
Db	411	CCGTACCAACACTCCGCGCGCAGGGCTACTCCGGTGGATGGACCTGTCTCTCAAGCCCG	470
Qy	581	TCGTGCGGGCTGCATCTGACCTCGATTCGAACCGGCGAGGAGCGGATCGGGAGGTGG	640
Db	471	TCGGGCTCAAGGCCACCCGCTACAGTCCGCGGAGCGGCAGAGAAGCTGCAGGAGC	530
Qy	641	CGGCGGCGCGCGAGCGCCGCA	662
Db	531	CCGGGTACGTCACTGTCGGCAA	552
RESULT 10			
CK626160			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			

RESULT 10				
CK626160				
LOCUS	CK626160	584 bp	mRNA	linear EST 26-JAN-2004
DEFINITION	mj18g02.y1 Mouse RPE/choroid, unamplified: mi/mj Mus musculus cDNA clone mj18g02 5', mRNA sequence.			

```

FEATURES
source
seq primer: MJ3KPL reverse primer (AB1).
Location/Qualifiers
1..584
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL6J"
/db_xref="taxon:10090"
/clone="mj18902"
/sex="Male"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMDH108"
/clone_lib="Mouse RPE/choroid, unamplified: mj/mj"
/note="Organ: Eye; Vector: pSport1; 64ug total RNA was
extracted from 200 adult male mouse RPE/choroids. A
directionally cloned cDNA library in the pSPORT1
vector (Life Technologies) was constructed at Bioserve
Biotechnology (Laurel MD) essentially following the

```

ORIGIN

Query Match	8.7%	Score 60.8	DB 7	Length 584
Best Local Similarity	47.4%	Pred. No. 0.019		
Matches 252	Conservative 0	Mismatches 272	Indels 8	Gaps 2
Qy	143	CGGTACCAAGGCGGATGTACTTCCACTTCGCTTCCAAAGGAAGAGCTGGCCCGCGGCG	202	
Db	61	CGCTGACGGAAGTCCGTGGTGGTGAACCCCCAGACGAAGACGCTCCCGTGTCTCGGCT	120	
Qy	203	TGCTGGCCGAGCAGACCCCTGCAGCTGGCGGTGCGGAAATCCGGCTTCCAAAGCGCAGGAAC	262	
Db	121	TCGTGACCTACTCGAACGTGGAGAGGCGGATGCGCCATGCGCGGTGCCGCACGCGG	180	
Qy	263	TGTTAGACTTCAACATGTGTGGTGCCTCCAGCGCATGTGTGACAGATCCGATCCTTCGGGGCGG	322	
Db	181	TGGACGGCAACACGGTGGAGCTGAAGCGCGCCGTGTGCGGGAGGATTCGCGCGCGGCCG	240	
Qy	323	GCACGGGGCTCGCACTGGACAGGGGGCGGTGACATTCCTCCGACGCCAACCCGTTTCGGCG	382	
Db	241	GGCGCGCACGCAAGGTGAAGAGCTGTGTGGCGGCTTCAAGGGCGAC--GTGCGG	297	
Qy	383	AGTGGGGCGACATCTGCGCCAGCTCTGTCGGCGAGGCACAGGAACCGGGGGAGGTGCTTC	442	
Db	298	AGGGGACCTGATCGAGCACTTCTCGAGTTCGGCGGGTGGAGAGCGGAGATCATTTG	357	
Qy	443	CGCACGTGAACCCGAAAAAGACCGGCGCATTTCAATCGTCGGCTCTTACCGGSGCTCAGG	502	
Db	358	CCGACAAGCAGTCCGGCAAGAAGCGCGCTTCGGCTTCGTCTACTTCCAGAGCCACGACG	417	
Qy	503	CGGTCTCCGGGTACCTTCGACCGCGCAGGACCTTCGGCCACCGGATCTCGTGTATGTGA	562	
Db	418	CGGCCGACACAGGCGCGGTGGTCAAGTTCACCCCGATCCAGGGCCACCGCGTGGAGGTGA	477	
Qy	563	ACCACGTGTGCCCGAGCATCTGTCGGCGGTCCATGTGACTCGATCGAAACCGGCGAGG	622	
Db	478	AGAAAGCGGTGCCCA-----AGAGAGATATCCACGCGGGCGCGGGGTGCGGGGGCGG	532	
Qy	623	AGCGGATTCGGGAAGGTGCGGGCGCGCGCGAGGCCCGCGAGGCTGCGGAGGC	674	
Db	533	CCGGGGCGGCGCGCGGAGGCTCGGGGCGCGCGCGGAGGCGCGCGGCGGCGGCGGCGG	584	

RESULT 11

LOCUS	CN011285	657 bp	mRNA	linear	EST 29-MAR-2004
DEFINITION	WHE3882_C08_F16ZS Wheat Fusarium graminearum infected spike cDNA library Triticum aestivum cDNA clone WHE3882_C08_F16, mRNA sequence.				

COMMENT

Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderan@pw.usta.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20. No effort was taken
to identify ESTs of fungal origin from this library, thus this EST
could be of wheat or fungal origin.

Seq primer: SK primer.

FEATURES

source

Location/Qualifiers
1..657
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Sumai3"
/db_xref="taxon:4565"
/clone="WHE3882 C08_F16"
/issue_type="Spike"
/tissue_type="Adult plant"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/clone_lib="Wheat Fusarium graminearum infected spike cDNA
library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid
pluascript SK; Site 1: EcoRI; Site 2: XhoI; Plants were
grown in the greenhouse. Spikes were sprayed at anthesis
with Fusarium graminearum. Total RNA, and poly(A) RNA were
prepared and pooled from infected spike at 0, 6, 12, 24,
36 and 48 hours after inoculation, a cDNA library was
made, and the cDNA clones were in vivo excised to give
pluascript phagemids in G. Muehlbauer lab at the
University of Minnesota (Kruger, W.M., Muehlbauer, G.J.,
Pritsch, C., Vance, C.). The cDNA library should contain
genes of both wheat and fungal pathogen origin. Plasmid
DNA preparations and DNA sequencing were performed in the
OD Anderson lab (all other authors)."

ORIGIN

Query Match 8.6%; Score 60; DB 7; Length 657;
Best Local Similarity 44.0%; Pred. No. 0.027;
Matches 255; Conservative 0; Mismatches 325; Indels 0; Gaps 0;

QY 63 GATCGTGGGGGAGCGCCCTCGGTCTTCGACGAGTACGGGTTCGAGCGCCACAGTGGC 122
DB 4 GCTCGTCCAGGCGGTCAAGCCGAGTATGCGCGCTGCGCGTGTCTCGCGGTGACCCG 63

QY 123 AGAGATCTCTCGCGGGCTCGGTACCAAGGGCGGATGTACTTCCATTCCTTCCAA 182
DB 64 CGACGTGCTCGCGCTCACCCTGCTCGACAAAGGCCAACTTCGACAGTACATCTCGGTGAT 123

QY 183 GGAAGAGCTGGCGCGGGCGGTGTGGCGGAGCAGACCTTGCACGTGGCGGTGCGGAATC 242
DB 124 CTACGCTTCCAGCACCGCGCGCGCCGCAACGACGTCTTCGAGGCGGGCTGCGCCAGGC 183

QY 243 CGGCTTCAAAGGCGCAGGAATCGTAGACCTTCAATGCTGGTGGTCCGCGGATGTCGA 302
DB 184 GCTGGTGGACTACCGCGAGTGGCGGGCGCTCGGTGTGGACGCCAACGGGACCGCGC 243

QY 303 CGATCGATCTTGGCGGGCGGGCGGCTCGCATCTGACAGGAGGGGGCGGTGACTTCTC 362
DB 244 GATCATGTCTCAACGACCGCGCGCGCGGTTCGTGGAGGCGACGCGCGAGCTCGCGCTCA 303

QY 363 CGACGCGCAACCCGCTTCGGCGAGTGGGGCGACATCTCGGCCCGAGCTCTCGCGCGAGGACA 422
DB 304 CAGCGTCATGCGCGCTTAAGCCCACTGTCGAGGTGCTGAACCTGACCCGAGCGGACGA 363

QY 423 GGAACGGGGGAGGTCTTTCGACAGTGAACCCGAAAAAGACCGCGACTTATCGTTCGG 482
DB 364 CGGACCGAGGAGTGTGATCTGATCCAGGTCAACGCGCTTCCCGTGGGGTCTGCTCGT 423

QY 483 CTGCTTCAAGGGCTCCAGGCGGTCTCCGGGTCACTCGGACCGCGAGCTTCGGCCA 542

DB 424 GGGGTTCCACGACGACACATCGTGTCCGACGGCGCTCCACCGGAACTTCTTCTCGTCC 483

QY 543 CGGATCTCGGTGATGTGAACACCGTGTGCCAGCATCGTGGCGGTCTCCATGCTGAC 602

DB 484 CTGGAGCAGGACCCCGGGCGCCCATCGACCCGCTCCAGTGCAGCAGCGGCTTC 543

QY 603 CTGGATCGAAACCGCGAGGAGCGGATCGGGAAGTCCGC 642

DB 544 CTCTCTTCATCCCGCGAACCGCTGCAGTTCGAGTACGAG 583

RESULT 12
CNS006XK/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 935)
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY Cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuoto Oosawa and
Aaron Mammeter in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..935
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N09"
/clone_lib="RPCI-98"
/note="end : 77"

ORIGIN
Query Match 8.6%; Score 60; DB 9; Length 935;
Best Local Similarity 27.5%; Pred. No. 0.027;
Matches 112; Conservative 106; Mismatches 190; Indels 0; Gaps 0;

QY 292 GGCATGTGCAGATCCGATCTCGGGGGGCGACGCGCTCGACTGACACGAGGGCG 351
DB 934 GSCGGGGGGGSGSGCGSGSGSGSCCGSCCGSCCGSCCGSCCGSCCGSCCGSC 875

QY 352 GTGGACTTCTCCGACGCCAACCCGTTCCGCGAGTGGGGGACATCTGCCCGCCAGTCTCTG 411
DB 874 SCGCGSGCGCCCGGSGSCCGCGSGSGSGCGCGCGCGCGCGCGCGCGCGCGCG 815

QY 412 GCGGAGGCGCAGGAACCGGGGGAGGTGCTTCCGACAGTGAACCCGAAAAAGACCGGCGAC 471
DB 814 GCGCGSGSCGCGCGCGSGCGSGCGSGSGSGCGSGSGSGCGSGSGSGCGSGSGSG 755

QY 472 TTTCATGTCGGTGTCTTACCGGGGCTCCAGGGGCTCTCCGGGTCACTTCGACCGCCAG 531


```

Db 754 GSCSSGSGSCGSGSGSCSSCGSCGSGSGSCGSCGSCGSCGSCGSCGSCGSG 695
Qy 532 GACCTCGGCGCACCGGATCTCGGTGATGTGGAACACAGTGTGCTGCCAGCATCTGTCGCGGCG 591
Db 694 CGSCGSGSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 635
Qy 592 TCCATCTGCTACCTGATCGMAACCGCGGAGGACCGATCGGGAAGTCTGCGGCGCGCGCG 651
Db 634 CCSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 675
Qy 652 GAGCGCGCGCGGAGCTCGGAGGCGCTCGGAGGCGCGCTCGGACGAGTAG 699
Db 574 SASGCGCGCGMGCRAGSGRMSAGSGRCGAGCGSGSGSGSSSRKGG 527

```

```

RESULT 13
LOCUS CK213336/c
DEFINITION FGAS025245 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum
aestivum cDNA, mRNA sequence.
CK213336
CK213336.1 GI:39619440

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 1084)
Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Geneswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
Penniket, C., Roach, J.L. and Sarhan, F.
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_estcs.usask.ca

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [20,670].
Plate: L6B008 row: P column: 08.
Location/Qualifiers
1. .1084
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"
/notes="Organ: Crown and leaf; Vector: pCMV.SPORTS; Crown
(50%) and leaf (50%) tissues from wheat cultivar Norstar
after short exposure times to low temperature in the light
and in the dark. 12 mRNA populations were combined before
constructing the library. The first 6 populations: After 7
days of growth at 20C8 from wheat cultivar Norstar after
short exposure times to low temperature in the light and
in the dark. 12 mRNA populations were combined before
constructing the library. The first 6 populations: After 7
days of growth at 20, wheat plants were transferred to 4C
in the light. 1cm crown sections and green leaf tissue were
separately harvested after 1, 3, and 6 hours of low
temperature exposure. The last 6 populations: After 7 days
of growth at 20C, wheat plants were transferred to 4C in
the dark. 1cm crown sections and green leaf tissue were
separately harvested after 1, 3, and 6 hours of low
temperature exposure. First strand synthesis in this

```

FEATURES

Source

library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI. In addition, this library used a primer for second strand synthesis that annealed to an artificial sequence (RNA oligo) added before first strand synthesis. Therefore when sequences from EST generated from this library will be masked for vector and adaptor sequences, an additional masking step will have to be included to mask this RNA oligo that is common to all clones (sequence CCACTGGAGCAGGACACTGACATGAGTGAAGGAGTAGAAA)."

ORIGIN

```

Query Match      8.6%; Score 59.8; DB 7; Length 1084;
Best Local Similarity 46.9%; Pred. No. 0.03;
Matches 184; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

```

```

Qy 279 GCTGGTCCGCGGCGGATGCTGCAGCATCCGATCTCTGCGGGGGGACGCGGCTCGCACT 338
Db 678 GGTCTTCAGCCAGTAGTCAAGGCGGAGCAGAGATGTCTCTCTCGCGGTTGTGCCCC 619
Qy 339 GGACCAAGGGGCGGTGGACTTCTTCGACGCCAACCCGTTTCGGCGAGTGGGGGACATCTG 398
Db 618 GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 559
Qy 399 GCGCCAGCTCTCTGGCGGAGGACAGGAACGGGGGAGGTGCTTCGCGACGTGAACCGGAA 458
Db 558 GGACCAAAACACGCGGCGCGCTGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 499
Qy 459 AAGACCGCGGCGCTTCATCGTCGGCTGTTTCCAGCGGCTCCAGGCGGTCTCCCGGCTCAC 518
Db 498 CGGCGGAGGTAGTGTCTTCGAGGACTTCCCGCGGACAGGCGCGCGAGGTCTATCGG 439
Qy 519 CTCGACCGCCAGGACCTCGGCGCACCGGATCTCGGTGATGTGGAACACAGTGTGCCCCAG 578
Db 438 CATGCCACACCGCGCGCGGTTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 379
Qy 579 CATCGTCCGCGGTCTCATCTGACCTGATCGAAACCGCGGAGGACGATCGGGAAGGT 638
Db 378 CATGGCAAGGCTCTCTCGAGAGGTTTTTCGAGAAGCGCAAGGACCGCTCGGGGCGG 319
Qy 639 CGCGGCGGCGCGGAGGCGCGCGGAGGCTCGG 670
Db 318 CGCGGCTTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 287

```

RESULT 14

AZ212924/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Trypanosoma brucei

Trypanosoma brucei

Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma

REFERENCE

AUTHORS

El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Suh, E., Malek, J., Fujii, C.,

Donelson, J., Frazer, C. and Adams, M.

Determination of clone end sequences from Trypanosoma brucei GUTat

10.1 sheared DNA library

Unpublished (1999)

CONTACT: Najib M. El-Sayed

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: nelsayed@igr.org

Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared

DNA library constructed at TIGR. Clones will be available for

distribution through Research Genetics, Alabama, USA. Sheared DNA
end sequences search page: <http://www.tigr.org/tdb/mdb/tbdb/>.
Seq primer: M13-Reverse
Class: shotgun.

FEATURES

Location/Qualifiers

1..483
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927/4 Gutat 10.1"
/db_xref="taxon:5691"
/clone="Sheared DNA-112H10"
/clone_lib="Sheared DNA"
/notes="Vector: pUC18; Site 1: SmaI; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 Gutat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. In Genome Sequencing: A Practical
Approach, eds. M. Vaudin and B. Borell, Oxford University
Press, 1999)."

ORIGIN

Query Match 8.5%; Score 59.2; DB 8; Length 483;
Best Local Similarity 49.5%; Pred. No. 0.04;
Matches 213; Conservative 0; Mismatches 208; Indels 9; Gaps 2;
QY 265 GTAGACCTCACCATGCTGCTGCCACGCGATGCTGCAGATCCGATCTGTGGGGGGGC 324
Db |||||
QY 454 GTAGGAGCAACCTGCTGTGCTCGCGACGCGCGGTGCAGCGGGGTGATCATCCGCTGGGC 395
QY 325 ACGCGGCTCGCATGTGACACGAGGGGGCGGTGGACTTCTCCGACGCCAACCCGTTCCGCGGAG 384
Db |||||
QY 394 AAGGGCTTCAACACGCTCAGGCCCTTGGCGACACCGTATCAAGCCGCGAGCCGTT 335
QY 385 TGGGGCGCATCTGCGCCCGAGTCTCTGCGGAGGACACAGGAACGCGGGGAGGTGTTCCG 444
Db |||||
QY 334 CCGACGCGCATCTGCGCCCGCAAGCGCGCGGCGGGGTGCGGGGTGGAATCTAT 275
QY 445 CACGTGAACCCGAAAGACCGCGACT-----TCATCGTGGTGTCTTACCGGGCTC 498
Db |||||
QY 274 GCGCGCATCCCGGACGATCGCGCGCGGTGATCATGAACGCGCGGTGCTACGGCTCG 215
QY 499 CAGGCGGTCTCCGGGTACCTCCGACCGCCA---GGACCTCGGCGCACCGGATCTCGGTG 555
Db |||||
QY 214 GAGACGTCATGTGTCAGTCTGTCGCGGTGATGATCGCGCGGTGTTGCGGAG 155
QY 556 ATGTGGAACCATGCTGCCAGCATCTGTGCGGGGTCCATGCTGACCTGGATCGAAACC 615
Db |||||
QY 154 CTGTGCTGAGGAGTGTGACCTACACCTACCGTCAAGCGGCTGCGAGGACGCGGAGACC 95
QY 616 GCGGAGGACGATCGGGAAGTTCGCGGGCGCGCGAGCGCGCGGCTCGGAGGCC 675
Db |||||
QY 94 GTATGTAATGACCGCATCTTCGAGGGGGCGCGGACGAGCCCGAGGCGATCAAGGCC 35
QY 676 TCCGAGGCGC 685
Db |||||
QY 34 CGCATGGCCG 25

RESULT 15
BE363278/c
LOCUS
DEFINITION WS1.61.D03.g1.A002 water-stressed 1 (WS1) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BE363278
VERSION BE363278.1 GI:9304835
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 640)

Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
Pratt,L.H.

An EST database from Sorghum: water-stressed plants

Unpublished (2000)

Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.

Seq primer: PolyTMix

High quality sequence start: 50

High quality sequence stop: 574

POLYA=No.

FEATURES

Location/Qualifiers

1..640

/organism="Sorghum bicolor"

/mol_type="mRNA"

/db_xref="taxon:4558"

/clone_lib="Water-stressed 1 (WS1)"

/note="Organ: Mix of 5-week old plants on days 7 & 8 after

water was withheld; Vector: Lambda Zap; Site 1: XhoI;

Site 2: EcoRI; The library was made from poly-A RNA in the

cloning vector lambda ZAP II. Clones to be sequenced were

prepared by mass excision."

ORIGIN

Query Match 8.4%; Score 59; DB 2; Length 640;
Best Local Similarity 49.3%; Pred. No. 0.044;
Matches 181; Conservative 0; Mismatches 185; Indels 1; Gaps 1;
QY 316 CGGCGGCGACGCGGCTCGCACTGGACACGAGGGGGCGGTGGACTTCTCCGACGCCAACCG 375
Db |||||
QY 368 CGCCGCTCTGCGCGCGCGCAGGTACACGTGGCGCGCACCCACCCCGCGCGCGCG 309
QY 376 TTGCGGAGTGGGGCGACATCTGCCCGCAGTCTCTGGCGGAGGACACAGAACGGGGGGAG 435
Db |||||
QY 308 CTTGCGGTGCGCCTCGTCTGCTGCCCATCTCTCAATCTCGAACAGCAGCGCGGCTC 249
QY 436 GTGCTTCCGACGTGAACCCGAAAGAACCGCGGACTTTCATCGTCGGCTGCTTACCGGG 495
Db |||||
QY 248 CGGTCCGGAGCGGGAAGCGGGCGGCGCACGCGGGAACCTCGACGGGCGCGACGCCGC 189
QY 496 CTCAGGCGGTCTCCCGGGTCACTTCCGACCGCCAGGACCTCGGCGACCGGATCTCGGTG 555
Db |||||
QY 188 CACTTGGCGCGCGCGAGCGCGTCCCGAACCGCTGGCTCAGCCCTCCCTGCACGCCGC 129
QY 556 ATGTGGAACCATGCTGCCCGCAGCATCG-TGCGGCGTCTCATGTGACCTGGATCGAAAC 614
Db |||||
QY 128 CCGCGGGCGCGCGCTCGCGCGCGGCGCTCTTGCCTTCCACTTCCACTTGGCGGGAGCG 69
QY 615 CGGCGAGGAGCGGATCGGGAAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674
Db |||||
QY 68 CGCGGGGAGGGGCTCCCGGGTTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 9
QY 675 CTCGAG 681
Db |||||
QY 8 CTCGCGC 2

Search completed: March 6, 2005, 22:12:04
Job time : 3474 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 20:07:07 ; Search time 149 Seconds
(without alignments)
602.204 Million cell updates/sec

Title: US-10-049-710A-2
Perfect score: 1164
Sequence: 1 MAVRHVAVRQRAVTRTQ.....VAAAEAAEAASEAASDE 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1164	100.0	232	5	AAM48990 S virgini
2	476	40.9	215	5	Abb80940 S. coelic
3	476	40.9	215	7	Adel14791 Streptomy
4	434.5	37.3	222	8	Adh39751 Streptomy
5	383	32.9	80	8	Adh39814 BarA amin
6	338.5	29.1	224	3	Ray83794 S. fradia
7	243	20.9	214	4	Aau04043 Streptoco
8	236	20.3	228	8	Adh39748 Streptomy
9	233	20.0	80	8	Adh39815 Fara amin
10	224	19.2	80	8	Adh39813 Arpa amin
11	214	18.4	213	8	Adh39743 Streptomy
12	161.5	13.9	203	4	Aau04039 Streptoco
13	138.5	11.9	211	6	Ada34808 Acinetoba
14	138	11.9	195	2	Aaw2428 Glucuroni
15	135.5	11.6	200	6	Abul6923 Protein e
16	134	11.5	201	7	Abos2144 Pseudomon
17	132.5	11.4	295	6	Abu21858 Protein e
18	131.5	11.3	273	6	Abu20218 Protein e
19	125	10.7	111	4	Abbs3089 Escherich
20	124.5	10.7	253	8	Adi39286 S. hygrois
21	124.5	10.7	313	4	Abg30006 Novel hum
22	122	10.5	295	5	Abba48122 Listeria
23	120.5	10.4	188	4	Aag91446 C glutami
24	120.5	10.4	188	7	Add13970 C. glutam
25	120.5	10.4	202	5	Abb49955 Listeria

26	120	10.3	205	6	ABU21324	Abu21324 Protein e
27	119	10.2	229	7	ABO63823	Abos3823 Klebsiell
28	118.5	10.2	81	6	ABU22851	Abu22851 Protein e
29	118.5	10.2	216	7	ABO71378	Abos71378 Pseudomon
30	118	10.1	280	6	ABR55197	AbR55197 Amino aci
31	117.5	10.1	231	6	ABM69166	Abm69166 Photorhab
32	115.5	9.9	189	4	ABBS2817	Abbs2817 Escherich
33	115.5	9.9	230	7	ABO67178	Abos7178 Klebsiell
34	115.5	9.9	234	7	ADF04427	Adf04427 Bacterial
35	115	9.9	182	6	ABU34193	Abu34193 Protein e
36	115	9.9	254	7	ABO69312	Abos69312 Pseudomon
37	114	9.8	194	6	ABU19797	Abu19797 Protein e
38	114	9.8	212	6	ABJ18817	Abj18817 Pseudomon
39	114	9.8	213	8	ADQ03074	Adq03074 P. aerugi
40	114	9.8	230	7	ABO73318	Abos7318 Pseudomon
41	113.5	9.8	213	7	ABO79293	Abos79293 Pseudomon
42	113	9.7	199	6	ADB12647	Adb12647 Alloiococ
43	113	9.7	214	6	ABM70661	Abm70661 Photorhab
44	113	9.7	243	7	ABO78019	Abos78019 Pseudomon
45	112.5	9.7	235	7	ABO67541	Abos67541 Klebsiell

ALIGNMENTS

RESULT 1
AAM48990
ID AAM48990 standard; protein; 232 AA.

XX AAM48990;

XX DT 03-MAY-2002 (first entry)

XX DE S virginiiae butanolide binding protein.

XX KW Butanolide binding protein; gene expression induction; operator;

XX KW transgenic plant; antibiotic production.

XX OS Streptomyces virginiiae.

XX PN WO200196581-A1.

XX PD 20-DEC-2001.

XX PF 15-JUN-2001; 2001WO-JP005096.

XX PR 15-JUN-2000; 2000JP-00180466.

XX PA (KANF) KANEKA CORP.

XX PI Shimmyo A, Kato K, Yamada Y, Nihira T, Shindo T;

XX DR WPI; 2002-098073/13.

XX DR N-PSDB; AAL44719.

XX PT Inducing expression of gene under regulation by operator at actinomycetes self-regulator provision site, useful in producing transformant tobacco for production of antibiotics e.g. virginiamycin.

XX PS Disclosure; Page 48-49; 57pp; Japanese.

XX CC The present invention relates to a method of inducing the expression of a gene under the regulation of an operator at an actinomycetes self-regulator provision site comprising imparting the characters of a repressor and an operator constituting a gene expression-inducing system with use of the actinomycetes self-regulator as inducer to a plant by gene transfer thereby providing the self-regulator to the transformed CC plant. The method is useful in producing transformant tobacco for CC production of antibiotics e.g. virginiamycin. The present sequence is the CC Streptomyces virginiiae butanolide binding protein which was used in the CC exemplification of the invention

XX SQ Sequence 232 AA;

Query Match 100.0%; Score 1164; DB 5; Length 232;
Best Local Similarity 100.0%; Pred. No. 3.4e-120;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVRHERVAVRQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTGKAMYPHFAS 60
DB 1 MAVRHERVAVRQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTGKAMYPHFAS 60

QY 61 KEELARGVLAEOQLHVAVPESGSKAQLVDLTMLVAHGMHLHDPILRAGTRRLALDQGVDF 120
DB 61 KEELARGVLAEOQLHVAVPESGSKAQLVDLTMLVAHGMHLHDPILRAGTRRLALDQGVDF 120

QY 121 SDANPFGWGNDICALLAEQAERGEVLPHVNPCKTGDFTVGCFTGLQAVSRVTSRQDLG 180
DB 121 SDANPFGWGNDICALLAEQAERGEVLPHVNPCKTGDFTVGCFTGLQAVSRVTSRQDLG 180

QY 181 HRISVMNHNVLPSIVPASMILTWIETGEERIGKVAAAEAAEAAEAAEAAEAAEAAE 232
DB 181 HRISVMNHNVLPSIVPASMILTWIETGEERIGKVAAAEAAEAAEAAEAAEAAEAAE 232

RESULT 2
ABB80940
ID ABB80940 standard; protein; 215 AA.
XX
AC ABB80940;
XX
DT 21-OCT-2002 (first entry)
XX
DE S. coelicolor Scbr protein.
XX
KW Antibiotic; bacterium; scbA; afsA; scbR; arpA; barA; actinorhodine; Act;
KW undecylprodigiosin; Red.
XX
OS Streptomyces coelicolor.
XX
PN CA2322241-A1.
XX
PD 23-APR-2002.
XX
PF 23-OCT-2000; 2000CA-02322241.
XX
PR 23-OCT-2000; 2000CA-02322241.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Takano E, Bibb M;
XX
DR WPI; 2002-501089/54.
DR N-PSDB; ABB86379.
XX
PT Modifying antibiotic-producing Streptomyces, to increase, or alter timing
PT of, antibiotic production, by deleting the scbA or scbR genes.
XX
PS Claim 20; Fig 9; 64pp; English.
XX
CC The invention provides a method for modifying an antibiotic-producing
CC strain of Streptomyces to increase production of antibiotics or to alter
CC the timing of antibiotic production. The modification is functional
CC deletion of the scbA gene of S. coelicolor, or its homologues, but is not
CC deletion of the afsA gene of S. griseus, or the modification is
CC functional deletion of the scbR gene of S. coelicolor, or its homologues,
CC but is not deletion of arpA of S. griseus nor barA of S. virginiae. The
CC method is particularly used for production of the antibiotics
CC actinorhodine (Act) and undecylprodigiosin (Red). The present sequence
CC represents the S. coelicolor Scbr protein

Sequence 215 AA;
Query Match 40.9%; Score 476; DB 5; Length 215;
Best Local Similarity 47.3%; Pred. No. 5.9e-44;
Matches 97; Conservative 36; Mismatches 72; Indels 0; Gaps 0;

QY 11 RQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTGKAMYFHFASKEELARGVLA 70
DB 3 KQDRAIRTRQITLDRAAQVFEKQGYQAATITILKAVGTYKALYFHFQSKELALGVDF 62

QY 71 EOTLHVAVPESGSKAQLVDLTMLVAHGMHLHDPILRAGTRRLALDQGVDFSDANPFGWG 130
DB 63 AQEPQAVPEQPLRLQELIDMGLFCHRLRTNVVARAGVRLSMDQQAHLDRGPFRRWH 122

QY 131 DICAQLLAQAERGEVLPHVNPCKTGDFTVGCFTGLQAVSRVTSRQDLGHRISVMNHN 190
DB 123 ETLLKLLNQAKENGELPHVVTDSADLYVGTGAGIQVVSQVTSYQDLHRYALLQKHI 182

QY 191 LPSIVPASMILTWIETGEERIGKVAA 215
DB 183 LPAIAPSVLAALDLSEERGAELAA 207

RESULT 3
ADE14791
ID ADE14791 standard; protein; 215 AA.
XX
AC ADE14791;
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptomyces coelicolor Scbr protein.
XX
KW antibiotic-producing strain; antibiotic production; scbA gene; Scbr gene;
KW actinorhodin; undecylprodigiosin.
XX
OS Streptomyces coelicolor.
XX
PN US2003124644-A1.
XX
PD 03-JUL-2003.
XX
PF 23-OCT-2001; 2001US-00017471.
XX
PR 23-OCT-2000; 2000US-0242561P.
XX
PA (TAKA/) TAKANO E.
PA (BIBB/) BIBB M J.
XX
PI Takano E, Bibb MJ;
XX
DR WPI; 2003-810983/76.
DR N-PSDB; ADE14790, ADE14794.
XX
PT Modifying an antibiotic-producing strain of Streptomyces coelicolor or
PT Streptomyces lividans to increase or alter the timing of antibiotic
PT production in the strain, comprises functionally deleting in the strain
PT the scbA or Scbr gene.
XX
PS Claim 28; Fig 9; 33pp; English.
XX
CC This invention relates to the novel modification of an antibiotic-
CC producing strain of Streptomyces coelicolor or Streptomyces lividans to
CC increase or to alter the timing of antibiotic production in the strain.
CC The method comprises functionally deleting in the strain the scbA or Scbr
CC gene. The method is useful in increasing and altering the timing of
CC antibiotic production (especially actinorhodin and undecylprodigiosin) in
CC Streptomyces species, particularly Streptomyces coelicolor or
CC Streptomyces lividans. The present sequence is that of the Streptomyces
CC coelicolor Scbr protein which is used in the method of the invention.

Sequence 215 AA;
Query Match 40.9%; Score 476; DB 7; Length 215;
Best Local Similarity 47.3%; Pred. No. 5.9e-44;
Matches 97; Conservative 36; Mismatches 72; Indels 0; Gaps 0;

QY 11 RQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTGKAMYFHFASKEELARGVLA 70

XX Streptomyces coelicolor.
OS WO200148228-A1.
PN XX
XX 05-JUL-2001.
PD XX
XX 20-DEC-2000; 2000WO-GB004972.
PF XX
XX 23-DEC-1999; 99GB-00030477.
PR XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
PA XX
XX Chater KF, Bruton CJ, O'rouke SJ, Wietzorrek AW;
PI WPI; 2001-425675/45.
XX N-PSDB; AAS07627, AAU04045, AAU04046.
DR XX
XX Novel expression cassette for expressing a nucleic acid of interest, PT
PT derived from the regulatory region of methylenomycin gene cluster of SCP1
PT plasmid of Streptomyces coelicolor A3(2).
XX
XX Claim 25; Fig 8a; 142pp; English.
PS
XX The sequence represents the MmyR protein encoded by the mmyR gene carried
CC on the expression cassette present on plasmid SCP1. The expression
CC cassette is the regulatory region of the methylenomycin cluster (mmc)
CC from Streptomyces coelicolor A3(2), which encodes the MmyR, MmfP, MmfH,
CC MmfL, MmfR, MmyT, MmyG, MmyJ and partial Mmr polypeptides. The
CC expression cassette is useful for expressing a nucleic acid of interest,
CC substantially only when the host cell culture reaches high cell density
CC at or close to the stationary phase of host cell culture. In particular
CC the system is useful in regulating methylenomycin production. Reduced or
CC no expression of the nucleic acid of interest is observed earlier in
CC growth, avoiding toxic effects of some gene products on growth and the
CC cluster naturally present on a highly transmissible plasmid permits
CC properly regulated expression in diverse Streptomyces host and the
CC expression is driven by a strong promoter, leading to high yield of the
CC desired end product. (Updated on 11-SEP-2003 to standardise OS field)
XX
XX Sequence 203 AA;
SQ
Query Match 13.9%; Score 161.5; DB 4; Length 203;
Best Local Similarity 34.1%; Pred. No. 4e-09;
Matches 44; Conservative 19; Mismatches 45; Indels 21; Gaps 3;
QY 10 VROERAVTRQAIIVRAAASVDEYGFEEATVAEILSRASVTKGMYFHFASKEELARGVL 69
Db 1 VQOARAVTRDQVLDAAAEFEALHGAGTNLATVAVRTGTMGALYGHFPFKKALADELV 60
QY 70 AEQT-----LHVAVPESGSKAQLVDITLVAHGMHLDPILRAGTRIALD---- 114
Db 61 SQSTETWNTIGRSIAETACAPETALRA-----LVLAVSRQMKHDIRFRAALRLAADCTMP 115
QY 115 -QGAVDFSD 122
Db 116 AGGAPDLLD 124
RESULT 13
ADA34808
ID ADA34808 standard; protein; 211 AA.
XX
XX ADA34808;
XX
XX 20-NOV-2003 (first entry)
DT
XX Acinetobacter baumannii protein #1969.
DE
XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX

OS Acinetobacter baumannii.
XX US562958-B1.
PN XX
XX 13-MAY-2003.
PD XX
XX 04-JUN-1999; 99US-00328352.
PF XX
XX 09-JUN-1998; 98US-0088701P.
PR XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Breton G, Bush D;
PI WPI; 2003-576092/54.
XX N-PSDB; ADA30682.
DR
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
XX Example; SEQ ID NO 6095; 328pp; English.
PS
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
XX Sequence 211 AA;
SQ
Query Match 11.9%; Score 138.5; DB 6; Length 211;
Best Local Similarity 22.9%; Pred. No. 1.5e-06;
Matches 49; Conservative 34; Mismatches 74; Indels 57; Gaps 5;
QY 1 MAVRHERVAVRQERAVRTRQAIIVRAAASVDEYGFEEATVAEILSRASVTKGMYFHFAS 60
Db 4 LIVVTDECKRLNKSETTRQIHLDTSFELVLHKGFGVGLQELIKACDVPKGSFYHYFAS 63
QY 61 KEELARGVL-----AEQTLHAV-----PESGSKAQE--LVDL 91
Db 64 KEAFGALLLEQYMWYKVRMEQLWQHSEQSAHARLMALWQAWIDDPVHGSWAENCLIVKL 123
QY 92 TMLVA-----HGMHLDPILRAGTRIALDQGVDFSDANPFGEWGDICALLAEQERGE 145
Db 124 AAEVSDLSSEDMRQILNDGVHKLTORLAL-----LUKEGQOEGS 161
QY 146 VLPHVNPKTGDFIVGCFGLQAVSRVTSDDROL 179
Db 162 IPKHIDPLKTAQVMYQLWLGALLIKLSQDKAHL 195
RESULT 14
AAW42428
ID AAW42428 standard; protein; 195 AA.
XX
XX AAW42428;
XX
XX 06-JUL-1998 (first entry)
DT
XX Glucuronide repressor GusR.
DE
XX Glucuronide repressor; GusR; transgene detection; assay.
KW Escherichia coli.
XX
XX Key Location/Qualifiers
FH Domain 1..63
FT

FT Domain /note= "operator binding domain"
FT 64. 154
FT /note= "glucuronide binding domain (may extend to residue
FT 195)"
PN WO9749813-A2.
XX 31-DEC-1997.
XX 26-JUN-1997; 97WO-US011427.
XX 26-JUN-1996; 96US-0020621P.
XX (CMB-) CAMBIA BIOSYSTEMS LLC.
XX Jefferson RA, Wilson KJ, Leader M;
PI WPI; 1998-077176/07.
XX N-PSDB; AAV03424, AAV03430.
DR Glucuronide repressor gene - useful to develop products to detect
PT glucuronide(s), control gene expression or identify glucuronide transport
PT proteins.
XX Claim 12; Page 40; 84pp; English.

XX This polypeptide nucleic acid molecule comprises the glucuronide
CC repressor (GusR) of Escherichia coli. The existence of the GusR gene (see
CC AAV03424) was established by genetic and biochemical experiments and
CC genetically mapped to a region upstream of the glucuronidase gene gusA
CC (see also AAV03430). GusR acts by binding to gusA operator sequences (see
CC AAV03429), thus preventing transcription, this repression being relieved
CC when a glucuronide substrate binds to the repressor and inactivates it.
CC Overexpression of the GusR gene product is achieved by cloning the coding
CC region in an expression vector. Vectors comprising a nucleic acid
CC molecule encoding GusR, its glucuronide binding domain or a fusion
CC protein, especially a binary Agrobacterium tumefaciens plasmid vector, as
CC well as plant, animal, fungal and bacterial host cell transformed by such
CC a vector are claimed. The repressor is used in claimed methods for
CC controlling expression of a transgene, detecting glucuronides in a sample
CC and isolating glucuronides from a sample

XX SQ Sequence 195 AA;

Query Match 11.9%; Score 138; DB 2; Length 195;
Best Local Similarity 26.3%; Pred. No. 1.5e-06;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;
QY 12 QERAVTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTYKGMYPFASKEELARG-VLA 70
DB 5 QTEAQPTTRILNAAAREIFSENGFHSASMKACKSCAISPGLYHHFISKEALIQAILQ 64
QY 71 EOTLHVA---VPESG-----SKAQELVDLT-----MLVAHGMHDPILRAGTRL 111
DB 65 DQERALARFREPIEGHIFVDYMWESIVSLTHEAFQGPALVETMAEGM-RNPQVAAMLK- 122
QY 112 ALDQGAVFDSANPFGSEWGDICQLLAAEQERGEVLPVHNPVKTG-----DFIVGCTGLQ 167
DB 123 -----NKHMTITEFVAQRMRDAQKGEISPDINTAMTSRLLLDLTYGVLAIE 170
QY 168 A---VSRVTSRQDL 179
DB 171 AEDLAREASPAQGL 184

RESULT 15
ABU16923
ID ABU16923 standard; protein; 200 AA.

XX ABU16923;
XX 19-JUN-2003 (first entry)
XX

DE Protein encoded by Prokaryotic essential gene #2450.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Acinetobacter baumannii.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ARA20793.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 44847; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 200 AA;

Query Match 11.6%; Score 135.5; DB 6; Length 200;
Best Local Similarity 23.5%; Pred. No. 3e-06;
Matches 48; Conservative 30; Mismatches 69; Indels 57; Gaps 5;

QY 11 RQERAVTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTYKGMYPFASKEELARGVL- 69
DB 3 RLNKSETTRQHLDTSLFELVHLKGFVGLQELKACDVPKGSFYHFAFASKEAFGCALLE 62

```

Qy 70 -----AEQTLHVAV-----PESGSKAQE--LVDLTMLVA----- 96
Db 63 OYMANYKVRMEQLWOHSEQSAHARLMAWQAWIDDPVHGSAENCLIVKLAEEVSDLS 122
Qy 97 -HGMHLDPILRACGTRIALDOGAVDFSDANPFGWGDICAOQLLAEOERGEVLPHVNP 155
Db 123 MRQILNDGVHKLTORLAL-----LLKEGQOEGSIPKHIEPLKT 160
Qy 156 GDFIVGCTGLQAVSRVTSRQDL 179
Db 161 AQVMYQLWLGAALLTKLSQDKAHL 184

```

Search completed: March 6, 2005, 22:29:13
Job time : 152 secs

This Page Blank (Uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 22:16:17 ; Search time 38 Seconds
(without alignments)
455.752 Million cell updates/sec

Title: US-10-049-710A-2
Perfect score: 1164
Sequence: 1 MAVRHERVAVRGERAVTRQ.....VAAAEAAAEASAEASDE 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A-COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B-COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A-COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B-COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCUTUS-COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138.5	11.9	211	4	US-09-328-352-6095
2	138	11.9	195	2	US-08-882-704A-2
3	138	11.9	195	4	US-09-151-957-2
4	134	11.5	201	4	US-09-252-991A-30890
5	133	11.4	190	3	US-09-303-120B-6
6	133	11.4	190	3	US-09-820-576-6
7	133	11.4	190	4	US-09-366-608-6
8	121	10.4	192	3	US-09-303-120B-8
9	121	10.4	192	3	US-09-820-576-8
10	121	10.4	192	4	US-09-366-608-8
11	119	10.2	229	4	US-09-489-039A-10340
12	118.5	10.2	216	4	US-09-252-991A-20124
13	115.5	9.9	230	4	US-09-489-039A-13695
14	115.5	9.9	234	4	US-09-543-681A-4712
15	115	9.9	254	4	US-09-252-991A-18058
16	114	9.8	230	4	US-09-252-991A-22064
17	113.5	9.8	213	4	US-09-252-991A-28039
18	113	9.7	210	3	US-09-199-637A-1172
19	113	9.7	243	4	US-09-252-991A-26765
20	112.5	9.7	235	4	US-09-489-039A-14058
21	109.5	9.4	207	4	US-09-328-352-5603
22	109	9.4	202	4	US-09-502-540-14870
23	108.5	9.3	198	4	US-09-376-451-5
24	108.5	9.3	223	4	US-09-543-681A-6126
25	108	9.3	204	4	US-09-134-000C-4883
26	108	9.3	206	4	US-09-328-352-6491
27	108	9.3	543	4	US-09-252-991A-27586

ALIGNMENTS

RESULT 1

US-09-328-352-6095
; Sequence 6095, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6095
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6095

Query Match 11.9%; Score 138.5; DB 4; Length 211;
Best Local Similarity 22.9%; Pred. No. 4e-08;
Matches 49; Conservative 34; Mismatches 74; Indels 57; Gaps 5;

QY 1 MAVRHERVAVRGERAVTRQAIIVRAAAVFDEYGFBAATVAETLSRAVTKGAMYHFAS 60
DB 4 LIVVYDECMKRLNKSETTRQHILDTSEFVLVHKGFVGVGLQEILKACDVPKSGFYHFAS 63
QY 61 KEELARGVL-----AQTLHVAV-----PESGSKAQB--LVDL 91
DB 64 KEAFGCALLEQYMANYSKVRMEQLWQHSQSAHRLMALWQAWITDDPVHGSWAENCLIVKL 123
QY 92 TWLVA-----HGLMHDPIRLAGTRIALDQGVDFSDANPFGEWGDICACOLLAAEQSRGE 145
DB 124 AAEVSDLSMDRQILNDGVHKLQRLAL-----LLKEGQGEQS 161
QY 146 VLPHVNPKITGDFIVGCFGLQAVSRVTSRDQDL 179
DB 162 IPKHIDPLTAQVMYQLWLGAALLTKLSQDKAHL 195

RESULT 2

US-08-882-704A-2
; Sequence 2, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; APPLICANT: Wilson, Katherine J.
; APPLICANT: Leader, Michael
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

ADDRESSES: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,704A
FILING DATE: 25-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5879906tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 190106.404
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear

US-08-882-704A-2

Query Match 11.9%; Score 138; DB 2; Length 195;
Best Local Similarity 26.3%; Pred. No. 4.1e-08;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;
QY 12 QERAVRTRQAIIVRAAASVFDEYGFEEATVAETLSRASVTKGAMYPHFASKEELARG-VLA 70
DB 5 QTEAQTTRILNAREIFSENGFHSKMAICKSAISPGTLYHHFISKEALIQAILIQ 64
QY 71 EQTLHVA---VPESG-----SKAQLVDLT-----MLVAHGMHDPILRAGTRL 111
DB 65 DQERALARFREPIEGIHFDYVYVESIVSLTHEAFQGRALVVEIMAEGM-RNQVAAMLK- 122
QY 112 ALDQGVDFSDANPFGWGDICALLAEAEQERGEVLPHVNPCKTG----DFIVGCFGTGLQ 167
DB 123 -----NKMTITFEVQMRDAQKQGEISPDINTAMTSRLLLDLTYGVLADE 170
QY 168 A--VSRVTSRDQDL 179
DB 171 AEDLAREASFAQGL 184

RESULT 3

US-09-151-957-2
Sequence 2, Application US/09151957
Patent No. 6429292
GENERAL INFORMATION:
APPLICANT: Jefferson, Richard A.
Leader, Michael
Wilsen, Katherine J.
TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,957
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/882,704
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 6429292tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 190106.404
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-151-957-2

Query Match 11.9%; Score 138; DB 4; Length 195;
Best Local Similarity 26.3%; Pred. No. 4.1e-08;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;
QY 12 QERAVRTRQAIIVRAAASVFDEYGFEEATVAETLSRASVTKGAMYPHFASKEELARG-VLA 70
DB 5 QTEAQTTRILNAREIFSENGFHSKMAICKSAISPGTLYHHFISKEALIQAILIQ 64
QY 71 EQTLHVA---VPESG-----SKAQLVDLT-----MLVAHGMHDPILRAGTRL 111
DB 65 DQERALARFREPIEGIHFDYVYVESIVSLTHEAFQGRALVVEIMAEGM-RNQVAAMLK- 122
QY 112 ALDQGVDFSDANPFGWGDICALLAEAEQERGEVLPHVNPCKTG----DFIVGCFGTGLQ 167
DB 123 -----NKMTITFEVQMRDAQKQGEISPDINTAMTSRLLLDLTYGVLADE 170
QY 168 A--VSRVTSRDQDL 179
DB 171 AEDLAREASFAQGL 184

RESULT 4

US-09-252-991A-30890
Sequence 30890, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30890
LENGTH: 201
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30890

Query Match 11.5%; Score 134; DB 4; Length 201;
Best Local Similarity 34.7%; Pred. No. 1.3e-07;
Matches 33; Conservative 15; Mismatches 35; Indels 12; Gaps 2;
QY 11 QERAVRTRQAIIVRAAASVFDEYGFEEATVAETLSRASVTKGAMYPHFASKEELARGVLA 70
DB 11 QQENAEATREALLESALSALFIEHGYGVSDAIAEARVTKGAFYHHFGSKQELLAECEYE 70


```
QY 71 EQTLHVA-----VP-----ESGSKAQLVLDLTM 93
Db 71 RQVRTIAEDLDVRPAHVWKWAEAAALAEAFIDSV 105

RESULT 5
US-09-303-120B-6
; Sequence 6, Application US/09303120B
; Patent No. 6309643
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBD-Associated Microbial Antigens and Methods of Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-PM 3478
; CURRENT APPLICATION NUMBER: US/09/303,120B
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Clostridium pasteurianum
US-09-303-120B-6

Query Match 11.4%; Score 133; DB 3; Length 190;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 49; Conservative 33; Mismatches 78; Indels 42; Gaps 7;

QY 17 RTRQAIVRAAASVFDEYGEFAATVAAILSRASVTYKGMATYFHFASKELARGVLAEE--QTL 74
Db 3 KTKDNIFYSAIKVFSNNGYNGATMDIASNAGVAKGTLYYHFKSKEEIKFYIIIEGVNLM 62

QY 75 HVAVPESGSKAQLVLDLTMVAHGMHLDPILRAGTTLALDQGAVDPSDANPFGEWGDICA 134
Db 63 KNEIDEATDKERTALEKLVACR-----VOLNLIYKNRDFFKVIASQLWKGKELR 111

QY 135 QL-----LAAEQERGEVLPHVNPVKTKGD--FIVGCFTG-LQAVSR---V 172
Db 112 QLELRDIRMNVYVHIEEFVKDAMEAGSI-----KKGNSLFVAYAFGLTCLCSVSLYEVI 164

QY 173 TSDRQDLGHRISVMNHNVLPSI 194
Db 165 NAENDNINNTIENLMNYILNGI 186

RESULT 6
US-09-820-576-6
; Sequence 6, Application US/09820576
; Patent No. 6320037
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid
; TITLE OF INVENTION: Molecules
; FILE REFERENCE: P-PM 4646
; CURRENT APPLICATION NUMBER: US/09/820,576
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/303,120
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Clostridium pasteurianum
US-09-820-576-6

Query Match 11.4%; Score 133; DB 3; Length 190;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 49; Conservative 33; Mismatches 78; Indels 42; Gaps 7;

QY 17 RTRQAIVRAAASVFDEYGEFAATVAAILSRASVTYKGMATYFHFASKELARGVLAEE--QTL 74
Db 3 KTKDNIFYSAIKVFSNNGYNGATMDIASNAGVAKGTLYYHFKSKEEIKFYIIIEGVNLM 62

QY 75 HVAVPESGSKAQLVLDLTMVAHGMHLDPILRAGTTLALDQGAVDPSDANPFGEWGDICA 134
Db 63 KNEIDEATDKERTALEKLVACR-----VOLNLIYKNRDFFKVIASQLWKGKELR 111

QY 135 QL-----LAAEQERGEVLPHVNPVKTKGD--FIVGCFTG-LQAVSR---V 172
Db 112 QLELRDIRMNVYVHIEEFVKDAMEAGSI-----KKGNSLFVAYAFGLTCLCSVSLYEVI 164

QY 173 TSDRQDLGHRISVMNHNVLPSI 194
Db 165 NAENDNINNTIENLMNYILNGI 186

RESULT 7
US-09-966-608-6
; Sequence 6, Application US/09966608
; Patent No. 6759530
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid
; TITLE OF INVENTION: Molecules
; FILE REFERENCE: P-PM 4966
; CURRENT APPLICATION NUMBER: US/09/966,608
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 09/303,120
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 09/820,576
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Clostridium pasteurianum
US-09-966-608-6

Query Match 11.4%; Score 133; DB 4; Length 190;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 49; Conservative 33; Mismatches 78; Indels 42; Gaps 7;

QY 17 RTRQAIVRAAASVFDEYGEFAATVAAILSRASVTYKGMATYFHFASKELARGVLAEE--QTL 74
Db 3 KTKDNIFYSAIKVFSNNGYNGATMDIASNAGVAKGTLYYHFKSKEEIKFYIIIEGVNLM 62

QY 75 HVAVPESGSKAQLVLDLTMVAHGMHLDPILRAGTTLALDQGAVDPSDANPFGEWGDICA 134
Db 63 KNEIDEATDKERTALEKLVACR-----VOLNLIYKNRDFFKVIASQLWKGKELR 111

QY 135 QL-----LAAEQERGEVLPHVNPVKTKGD--FIVGCFTG-LQAVSR---V 172
Db 112 QLELRDIRMNVYVHIEEFVKDAMEAGSI-----KKGNSLFVAYAFGLTCLCSVSLYEVI 164

QY 173 TSDRQDLGHRISVMNHNVLPSI 194
Db 165 NAENDNINNTIENLMNYILNGI 186

RESULT 8
US-09-303-120B-8
; Sequence 8, Application US/09303120B
; Patent No. 6309643
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBD-Associated Microbial Antigens and Methods of Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-PM 3478
; CURRENT APPLICATION NUMBER: US/09/303,120B
```

```

; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 192
; TYPE: PRF
; ORGANISM: Aquifex aeolicus
US-09-303-120B-8

Query Match          10.4%; Score 121; DB 3; Length 192;
Best Local Similarity 24.7%; Pred.No. 4.3e-06;
Matches 41; Conservative 31; Mismatches 46; Indels 48; Gaps 7;

QY      13 ERVVRTRQAIVRAAASVFDEYGFEATVAAILSRASVTKGAMYPHFASKELARGVLAEQ 72
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      9 EKRSDTKEKILSSALKLFSKKGPKETTKIDIAKEVGITEGAIVRHHFTSKBEIIKSLLESI 68
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      73 T-----LHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGAVDFSDANP- 125
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      69 TKELRHKLVAL-QRGTDDBEILE-----SIVDTLIDYAFSNPE 106
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      126 ---FGWGDICAOQLLAEAQRGEV--LPNVNPKKTGDFIVGCFTGL 166
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      107 SFRF-----LNLYHLLKEYGEVKNLFP-----GELLIKFLNGL 138
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 9
US-09-820-576-8
; Sequence 8, Application US/09820576
; Patent No. 6320037
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid
; FILE REFERENCE: P-PM 4646
; CURRENT APPLICATION NUMBER: US/09/820,576
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/303,120
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 192
; TYPE: PRF
; ORGANISM: Aquifex aeolicus
US-09-820-576-8

Query Match          10.4%; Score 121; DB 3; Length 192;
Best Local Similarity 24.7%; Pred.No. 4.3e-06;
Matches 41; Conservative 31; Mismatches 46; Indels 48; Gaps 7;

QY      13 ERVVRTRQAIVRAAASVFDEYGFEATVAAILSRASVTKGAMYPHFASKELARGVLAEQ 72
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      9 EKRSDTKEKILSSALKLFSKKGPKETTKIDIAKEVGITEGAIVRHHFTSKBEIIKSLLESI 68
       |||:::||:::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      73 T-----LHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGAVDFSDANP- 125
       |||:::||~::~||~::||~::||~::||~::||~::||~::||~::||~::||~::||~::||
Db      69 TKELRHKLVAL-QRGTDDBEILE-----SIVDTLIDYAFSNPE 106
       |||:::||~::~||~::||~::||~::||~::||~::||~::||~::||~::||~::||~::||

QY      126 ---FGWGDICAOQLLAEAQRGEV--LPNVNPKKTGDFIVGCFTGL 166
       |||:::||~::~||~::||~::||~::||~::||~::||~::||~::||~::||~::||~::||
Db      107 SFRF-----LNLYHLLKEYGEVKNLFP-----GELLIKFLNGL 138
       |||:::||~::~||~::||~::||~::||~::||~::||~::||~::||~::||~::||~::||

RESULT 10
US-09-966-608-8
; Sequence 8, Application US/09966608
; Patent No. 6759530
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid

```

RESULT 12
US-09-252-991A-20124
; Sequence 20124, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20124
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20124

Query Match 10.2%; Score 118.5; DB 4; Length 216;
Best Local Similarity 24.5%; Pred. No. 1e-05; Indels 53; Gaps 8;
Matches 47; Conservative 25; Mismatches 67; Indels 53; Gaps 8;

QY 12 QERAVR---TROAIVRAAASVFDEYCFEAAATVAEILSRASVTKGAMYFHFASKEELARG 67
Db 27 QEPSVRPLKISRDQLQRCAGTRRYGYHCTTMDLSSACGLTKASFYHHYPNKEALLRD 86
QY 68 VLAEOQLHVAVPESGSKAQELVDLTMLVAHGMHDPIL-----RAGTTLALDQG- 116
Db 87 VL---EWTHT-----QRLAETLFSIA-----YDPLLTPTRELKLGKRAARLFQDDSI 130
QY 117 -----AVDFSDANP-----FGWGDCIAQLAQAQERGEVLPHVNPCKTGFIV 160
Db 131 GCLMGVVAVDASVGRSELMAPIRSFLDDWAQAFQLYRPAFDEAQAAL-----ERGRQLV 184
QY 161 GCFTGLQAVSRV 172
Db 185 ADFEGAILLARI 196

RESULT 13
US-09-489-039A-13695
; Sequence 13695, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13695
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13695

Query Match 9.9%; Score 115.5; DB 4; Length 230;
Best Local Similarity 30.0%; Pred. No. 2.6e-05; Indels 27; Gaps 5;
Matches 36; Conservative 20; Mismatches 37; Indels 27; Gaps 5;

QY 12 QERAVRTRQAIIVRAAASVFDEYCFEAAATVAEILSRASVTKGAMYFHFASKEELARGVLAE 71
Db 20 KERAQRTQLLIESAIOQFALRGVNTTLTDIADAAGVTRGAVYWHFASKTELFENWQQ 79
QY 72 QTLHVAVPESGSKAQELVDLTMLVAHGMHDPIL--LR-----ACTRLALDQGVDFSDANP 125

Db 80 Q-----PPLRLDIQPSQAIEYE--HEPLNALRERFIAGLR-----YTAANP 118

RESULT 14
US-09-543-681A-4712
; Sequence 4712, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4712
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4712

Query Match 9.9%; Score 115.5; DB 4; Length 234;
Best Local Similarity 30.3%; Pred. No. 2.7e-05; Indels 13; Gaps 3;
Matches 40; Conservative 17; Mismatches 62; Indels 13; Gaps 3;

QY 11 ROERAVRTRQAIIVRAAASVFDEYCFEAAATVAEILSRASVTKGAMYFHFASKEELARGVLA 70
Db 34 QENSEQTRLALLEAQYLQVNCYYDVSDISRLRYARVTKGAFYHHSKALLRECYL 93
QY 71 EQLTHVA-----VPESGSKAQELVDLTMLVAHGMHDPILRAGTTLALDQGVDFSDANP 125
Db 94 LQVGHAVQKLDVPTVDDKQQLTALFSLCV-----DHIYQCKNELIPLQEVISVLG--- 145
QY 126 FGWGDICAQLL 137
Db 146 WKWEIBEIDANIL 157

RESULT 15
US-09-252-991A-18058
; Sequence 18058, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18058
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18058

Query Match 9.9%; Score 115; DB 4; Length 254;
Best Local Similarity 25.6%; Pred. No. 3.5e-05; Indels 10; Gaps 4;
Matches 40; Conservative 30; Mismatches 76; Indels 10; Gaps 4;

QY 13 ERAVRTRQAIIVRAAASVFDEYCFEAAATVAEILSRASVTKGAMYFHFASKEELARGVLAEQ 72
Db 51 ERGRQRRRLAALDAQAFLEHGFGFTTLDWIERAGSGRTLYSSFGKEGLFAAVIAHM 110
QY 73 QTLHVAVPESGSKAQELVDLTMLVAH-----GMLHDPILRAGTTLALDQGVDFSDANP 126

Db 111 -IEEIFDDSDQPRPAATLSATLEHFGRRFLTSLDPRCQSLYRLVVAESPRFFAICKSF 169

Qy 127 GEWDICQAQLLAQAQGEVLPVHNPXKKTGDFIVGC 162

Db 170 YEQGPQSQSYLL-SERLAAPVPHMDEETL--YAVAC 202

Search completed: March 6, 2005, 22:35:27
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: March 6, 2005, 22:27:08 ; Search time 128 Seconds
(without alignments)
596.394 Million cell updates/sec

Title: US-10-049-710A-2

Perfect score: 1164

Sequence: 1 MAVRHERVAVRQERAVRTRQ.....VAAAEAAAEAAAEASAEASDE 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTU5_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1164	100.0	232	14	US-10-049-710A-2
2	1159	99.6	232	15	US-10-168-663-9
3	495	42.5	221	15	US-10-168-663-7
4	476	40.9	215	14	US-10-017-471A-16
5	472	40.5	215	15	US-10-168-663-8
6	470.5	40.4	276	15	US-10-168-663-10
7	439.5	37.8	234	14	US-10-156-761-11242
8	272	23.4	215	15	US-10-168-663-5
9	257	22.1	215	15	US-10-168-663-6
10	248.5	21.3	218	14	US-10-156-761-11239
11	243	20.9	214	15	US-10-168-663-11
12	224	19.2	196	15	US-10-168-663-4
13	217.5	18.7	216	15	US-10-168-663-3

14	215.5	18.5	218	14	US-10-156-761-9807	Sequence 9807, Ap
15	207.5	17.8	287	14	US-10-156-761-11240	Sequence 11240, A
16	171.5	14.7	206	14	US-10-156-761-9805	Sequence 9805, Ap
17	161.5	13.9	203	15	US-10-168-663-12	Sequence 12, Appl
18	138	11.9	195	14	US-10-195-518-2	Sequence 2, Appl
19	138	11.9	227	14	US-10-156-761-13631	Sequence 13631, A
20	135.5	11.6	200	15	US-10-282-122A-44847	Sequence 44847, A
21	133	11.4	190	9	US-09-820-576-6	Sequence 6, Appl
22	133	11.4	190	9	US-09-966-608-6	Sequence 6, Appl
23	133	11.4	190	16	US-10-835-914-6	Sequence 6, Appl
24	132.5	11.4	295	15	US-10-282-122A-49782	Sequence 49782, A
25	131.5	11.3	273	15	US-10-282-122A-48142	Sequence 48142, A
26	131	11.3	199	14	US-10-156-761-11396	Sequence 11396, A
27	129.5	11.1	211	14	US-10-156-761-8108	Sequence 8108, Ap
28	127	10.9	197	14	US-10-156-761-12316	Sequence 12316, A
29	125	10.7	111	14	US-10-238-075-1557	Sequence 1557, Ap
30	124.5	10.7	253	15	US-10-461-194-128	Sequence 128, App
31	122.5	10.5	214	14	US-10-156-761-7977	Sequence 7977, Ap
32	121	10.4	192	9	US-09-820-576-8	Sequence 8, Appl
33	121	10.4	192	9	US-09-966-608-8	Sequence 8, Appl
34	121	10.4	192	16	US-10-835-914-8	Sequence 8, Appl
35	120.5	10.4	188	9	US-09-738-626-5200	Sequence 5200, Ap
36	120.5	10.4	188	17	US-10-495-066-36	Sequence 36, Appl
37	120	10.3	205	15	US-10-282-122A-49248	Sequence 49248, A
38	119.5	10.3	219	14	US-10-156-761-13447	Sequence 13447, A
39	119	10.2	194	14	US-10-156-761-8051	Sequence 8051, Ap
40	118.5	10.2	81	15	US-10-282-122A-50775	Sequence 50775, A
41	117.5	10.1	211	14	US-10-156-761-8127	Sequence 8127, Ap
42	117	10.1	220	14	US-10-156-761-10827	Sequence 10827, A
43	117	10.1	221	14	US-10-156-761-9317	Sequence 9317, Ap
44	115.5	9.9	189	14	US-10-238-075-1037	Sequence 1037, Ap
45	115.5	9.9	191	14	US-10-156-761-12235	Sequence 12235, A

ALIGNMENTS

RESULT 1

```

US-10-049-710A-2
; Sequence 2, Application US/10049710A
; Publication No. US20030126648A1
; GENERAL INFORMATION:
; APPLICANT: Shimmyo, Atsuhiko
; APPLICANT: Kato, Kou
; APPLICANT: Yamada, Yasuhiro
; APPLICANT: Nihira, Takuya
; APPLICANT: Shindo, Takuya
; TITLE OF INVENTION: METHOD FOR INDUCTION OF GENE EXPRESSION IN PLANT AND PLANT TREAT
; THEREBY
; FILE REFERENCE: 5405/18
; CURRENT APPLICATION NUMBER: US/10/049,710A
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: PCT/JP01/05096
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: JP 2000-180466
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Streptomyces virginiae
US-10-049-710A-2

```

Query Match 100.0%; Score 1164; DB 14; Length 232;

Best Local Similarity 100.0%; Pred. No. 5.9e-113; Mismatches 0; Indels 0; Gaps 0;

1 MAVRHERVAVRQERAVRTRQAIIVRAAASVDFYGFEEATVAEILSRASVTKGMYHFAS 60

1 MAVRHERVAVRQERAVRTRQAIIVRAAASVDFYGFEEATVAEILSRASVTKGMYHFAS 60

61 KEELARGVLAEGTLHVAVPESGSKAQELVDLTMLVHAGMLHDPILRAGTRRLDQGVDF 120

Db 61 KEELARGVLAETLHVAVPESGSKAQLVDLTMLVAHGMLHDPILRAGTRLALDQGVDF 120
QY 121 SDANPFGWGDDICAOALLAEQERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLG 180
Db 121 SDANPFGWGDDICAOALLAEQERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLG 180
QY 181 HRISVNMNHLPSIVPASMLTWETGEERIGKVAIAAAAEAAAEAAAEAAASDE 232
Db 181 HRISVNMNHLPSIVPASMLTWETGEERIGKVAIAAAAEAAAEAAAEAAASDE 232

RESULT 2

US-10-168-663-9
; Sequence 9, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909USO
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-10-168-663-9

Query Match 99.6%; Score 1159; DB 15; Length 232;
Best Local Similarity 99.6%; Pred. No. 1.9e-112;
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAVRHVAVRGERAVRTQAIIVRAAASVDFEYGFEEATVAAILSRASVTGKAMYPHFAS 60
Db 1 MAVRHVAVRGERAVRTQAIIVRAAASVDFEYGFEEATVAAILSRASVTGKAMYPHFAS 60
QY 61 KEELARGVLAETLHVAVPESGSKAQLVDLTMLVAHGMLHDPILRAGTRLALDQGVDF 120
Db 61 KEELARGVLAETLHVAVPESGSKAQLVDLTMLVAHGMLHDPILRAGTRLALDQGVDF 120
QY 121 SDANPFGWGDDICAOALLAEQERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLG 180
Db 121 SDANPFGWGDDICAOALLAEQERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLG 180
QY 181 HRISVNMNHLPSIVPASMLTWETGEERIGKVAIAAAAEAAAEAAAEAAASDE 232
Db 181 HRISVNMNHLPSIVPASMLTWETGEERIGKVAIAAAAEAAAEAAAEAAASDE 232

RESULT 3

US-10-168-663-7
; Sequence 7, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909USO
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25

; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-10-168-663-7

Query Match 42.5%; Score 495; DB 15; Length 221;
Best Local Similarity 46.8%; Pred. No. 3.7e-43;
Matches 101; Conservative 40; Mismatches 75; Indels 0; Gaps 0;
QY 12 QERAVRTQAIIVRAAASVDFEYGFEEATVAAILSRASVTGKAMYPHFASKEELARGVLA 71
Db 4 QVRAIRTRQAILSAARVDFEYGFEEATVAAILSRASVTGKALYFHFOSKEDLAGVLT 63
QY 72 QTLHVAVPESGSKAQLVDLTMLVAHGMLHDPILRAGTRLALDQGVDFSDANPFGWG 131
Db 64 QNEDLLLPKAEVVDVAVMLHTRLTNPMPVAGVRLSLDYNAGGLDRSAPFRNWD 123
QY 132 ICAQLLAERQERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLGHRISVNMNHL 191
Db 124 KFTDLLEKAAQAGELLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLGHRISVNMNHL 183
QY 192 PSIVPASMILTWTETGEERIGKVAIAAAAEAAAEAAAEAAEASE 227
Db 184 PSIAQPSVLAHLHGESRAEEVLEARQLAREQADE 219

RESULT 4

US-10-017-471A-16
; Sequence 16, Application US/10017471A
; Publication No. US20030124644A1
; GENERAL INFORMATION:
; APPLICANT: Takano, Eriko
; APPLICANT: Bibb, Mervyn
; TITLE OF INVENTION: Antibiotic Production
; FILE REFERENCE: 0380-P02329US1
; CURRENT APPLICATION NUMBER: US/10/017,471A
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,561
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-10-017-471A-16

Query Match 40.9%; Score 476; DB 14; Length 215;
Best Local Similarity 47.3%; Pred. No. 3.4e-41;
Matches 97; Conservative 36; Mismatches 72; Indels 0; Gaps 0;
QY 11 ROERAVRTQAIIVRAAASVDFEYGFEEATVAAILSRASVTGKAMYPHFASKEELARGVLA 70
Db 3 KODRAIRTRQITLIDAAQVFEKQGYQAATITELKLVAGVTGKALYFHFQSKELALGVDF 62
QY 71 EQTLHVAVPESGSKAQLVDLTMLVAHGMLHDPILRAGTRLALDQGVDFSDANPFGWG 130
Db 63 AQEPQAVPEQPLRLQELIDMGMLFCHRLRTNPVAVAGVRLSMDQAAHGLDRRGPFRWH 122
QY 131 DICAQLLAERQERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLGHRISVNMNHL 190
Db 123 ETLKLLNQAQKENGELLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLGHRISVNMNHL 182
QY 191 LPSIVPASMILTWTETGEERIGKVAIA 215
Db 183 LPAIAVPSVLAALDILSEEGARLAA 207

US-10-168-663-10

```

RESULT 5
US-10-168-663-8
; Sequence 8, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909USO
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25
; PRIORITY APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-10-168-663-8

```

Query Match	40.5%	Score 472;	DB 15;	Length 215;
Best Local Similarity	47.3%;	Pred. No. 8.9e-41;		
Matches 97;	Conservative 35;	Mismatches 73;	Indels 0;	Gaps 0
QY	11	RQERAVTRQAI	VRAAASVFDYGEFAATVA	ILSRASVTKGAMVFHFASKEELARGVLA 70
Db	3	KQDRAIR	QRTILDAQVFEKQYQAA	TIETLLKVAGVTKGALVFHFQSKKEELALGVFD 62
QY	71	EQTLLHAVPSSGS	KAQELAVDLTMLVAHGMLHD	PILRAGTRLALDQGVDFSDANPFGSWG 130
Db	63	ACEPPQAVPEOPL	LOELIDMGMLFCHRLRTN	VVARAGVRLSMDQQAAGLDRRGFFRW 122
QY	131	DICAAQLAAEQER	GEVLPHVNP	PKTGDPFVCGFTGLQAVSVTSDRDLGHRISVMMNHV 190
Db	123	ETLLFLNLQAKENG	ELPHVTTDSADLYVGT	FTFAGIQVVSQTVSDYQDLEHRYALLQKHI 182
QY	191	LPSIVPASMLT	WIETGEE	RIGKVA 215
Db	183	LPAAIVPSVLAAL	DLSEER	GARLAA 207

RESULT, T 7

```

US-10-156-761-11242
; Sequence 11242, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11242
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11242

```

Query Match	37.8%	Score 439.5;	DB 14;	Length 234;		
Best Local Similarity	42.6%;	Pred. No. 2.5e-37;				
Matches	95;	Conservative 44;	Mismatches 77;	Indels 7; Gaps 4;		
Qy	11	ROERAVRTRQAI	VRAAASVDFEYGF	FAATVAEILSRASVT	KGAMYHFHFAKSELARGVLA	70
Db	3	ROERAIRTRQ	ITLVAAREVDFE	VGYEATISDVLK	SGVTKGALYHFHTSKQELAAQVLA	62
Qy	71	EQTLLHV-AP	PESGSKAQELV	DLTMLVAHMLH---	DPILRAGTRTALADQAG--	124
Db	63	EQVASLPRV	PEQELKQQLS	DEALLAHLRECTG	DPVQGSVRLTVDQSGPRDHLNRRV	122
Qy	125	PFGEWGDI	CAQLLAAEAGE	RGEVLPHVNP	PKTGDFIVGCFTGLQASVRVTS	184
Db	123	PMQAWTEHT	QSLFEEBAR	KGELIPLHADV	EALAKLVGAFTGVQVLRIMTGRADLAERVA	182
Qy	185	VMNHNVLPS	IVPASM	LTWTGTEER	IGKV-AAAAEAAEAAAS	226
Db	183	DLYRHLMPS	FAMPGIILV	LDPSPERGSR	VVEAAKORSEAAAS	225

RESULT 6
US-10-168-663-10
; Sequence 10, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909U50
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 276
; TYPE: prt
; ORGANISM: Streptomyces sp.

Sequence 5, Application US/10168663
Publication No. US20040086962A1
GENERAL INFORMATION:
APPLICANT: Plant Bioscience Limited
APPLICANT: Chater, Keith F
APPLICANT: Bruton, Celis J
APPLICANT: O'Rourke, Sean J
APPLICANT: Wietzorrek, Andreas W
TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
FILE REFERENCE: 0380-P02909USO
CURRENT APPLICATION NUMBER: US/10/168,663
CURRENT FILING DATE: 2002-10-25
PRIOR APPLICATION NUMBER: PCT/GB00/04972
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: GB 9930477.6
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 215
TYPE: PRT
ORGANISM: Streptomyces sp.
US-10-168-663-5

Query Match 23.4%; Score 272; DB 15; Length 215;
Best Local Similarity 35.1%; Pred. No. 6.7e-20;
Matches 68; Conservative 35; Mismatches 85; Indels 6; Gaps 2;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTGKAMYFHFASKEELARGVLA 70
Db 3 RQRAERTRATVRAAADLFRGYESVTSLSIVAHAGVTGKALYFHFASKEELAHAIL 62
QY 71 EQ---TLHVAVPESG---SKAQELVDLTMLVAHGMHLDPIILRAGTRLALDQGVDFSDAN 124
Db 63 IQSRTSRRLAKDLDRGYSSLEALMRLTFGMARLCVQGVPLRAGRLATAGVVPVPLPH 122
QY 125 PFGEWGDICQAQLLAERAEQGEVLPHVNPCKTGDFTVGCFTGLQAVSRVTSRQDLGHRIS 184
Db 123 PFTDWREIATSRLLDAVRQSDVHQDIDVDSVAHTLVCSVVGTVVGGTLEPAGRPRLA 182
QY 185 VMNHNVLPSIVPAS 198
Db 183 EMWYILIRGMVPT 196

RESULT 10
US-10-156-761-11239
Sequence 11239, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11239
LENGTH: 218
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-11239

Query Match 21.3%; Score 248.5; DB 14; Length 218;
Best Local Similarity 33.3%; Pred. No. 1.9e-17;
Matches 68; Conservative 35; Mismatches 84; Indels 17; Gaps 5;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTGKAMYFHFASKEELAR--- 66
Db 3 RQRAERTRATVRAAADLFRGYESVTSLSIVAHAGVTGKALYFHFASKEELAHAIL 62
QY 67 --GVLAEQTLHVAVPESGSKAQELVDLTMLVAHGMHLDPIILRAGTRLALDQ---GAVDFS 121
Db 63 AAGVSLRRAAWLAQPPGTNALQRLTNTSHALERLGDVVARAGPRLNCSAGGGAL--- 119
QY 122 DANPFGEWGDICQAQLLAERAEQGEVLPHVNPCKTGDFTVGCFTGLQAVSRVTSRQDLG 180
Db 120 --NLLREWQTCVEQLLAERAEGLIARLVRADTVSAVVAATGTFELLOR--RDPWLSG 175
QY 181 HRISVMNHNVLPSIVPASMLTWIE 204
Db 176 QSLAAFWRVLLPRAATAAALTAVD 199

RESULT 11
US-10-168-663-11
Sequence 11, Application US/10168663
Publication No. US20040086962A1
GENERAL INFORMATION:
APPLICANT: Plant Bioscience Limited

Sequence 5, Application US/10168663
Publication No. US20040086962A1
GENERAL INFORMATION:
APPLICANT: Plant Bioscience Limited
APPLICANT: Chater, Keith F
APPLICANT: Bruton, Celis J
APPLICANT: O'Rourke, Sean J
APPLICANT: Wietzorrek, Andreas W
TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
FILE REFERENCE: 0380-P02909USO
CURRENT APPLICATION NUMBER: US/10/168,663
CURRENT FILING DATE: 2002-10-25
PRIOR APPLICATION NUMBER: PCT/GB00/04972
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: GB 9930477.6
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 215
TYPE: PRT
ORGANISM: Streptomyces sp.
US-10-168-663-5

Query Match 23.4%; Score 272; DB 15; Length 215;
Best Local Similarity 35.1%; Pred. No. 6.7e-20;
Matches 68; Conservative 35; Mismatches 85; Indels 6; Gaps 2;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTGKAMYFHFASKEELARGVLA 70
Db 3 RQRAERTRATVRAAADLFRGYESVTSLSIVAHAGVTGKALYFHFASKEELAHAIL 62
QY 71 EQ---TLHVAVPESG---SKAQELVDLTMLVAHGMHLDPIILRAGTRLALDQGVDFSDAN 124
Db 63 IQSRTSRRLAKDLDRGYSSLEALMRLTFGMARLCVQGVPLRAGRLATAGVVPVPLPH 122
QY 125 PFGEWGDICQAQLLAERAEQGEVLPHVNPCKTGDFTVGCFTGLQAVSRVTSRQDLGHRIS 184
Db 123 PFTDWREIATSRLLDAVRQSDVHQDIDVDSVAHTLVCSVVGTVVGGTLEPAGRPRLA 182
QY 185 VMNHNVLPSIVPAS 198
Db 183 EMWYILIRGMVPT 196

RESULT 9
US-10-168-663-6
Sequence 6, Application US/10168663
Publication No. US20040086962A1
GENERAL INFORMATION:
APPLICANT: Plant Bioscience Limited
APPLICANT: Chater, Keith F
APPLICANT: Bruton, Celis J
APPLICANT: O'Rourke, Sean J
APPLICANT: Wietzorrek, Andreas W
TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
FILE REFERENCE: 0380-P02909USO
CURRENT APPLICATION NUMBER: US/10/168,663
CURRENT FILING DATE: 2002-10-25
PRIOR APPLICATION NUMBER: PCT/GB00/04972
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: GB 9930477.6
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 215
TYPE: PRT
ORGANISM: Streptomyces sp.
US-10-168-663-6

Query Match 22.1%; Score 257; DB 15; Length 215;
Best Local Similarity 34.0%; Pred. No. 2.5e-18;


```

; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9807
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9807

```

```

Query Match      18.5%; Score 215.5; DB 14; Length 218;
Best Local Similarity 29.1%; Pred. No. 5.4e-14;
Matches 57; Conservative 39; Mismatches 85; Indels 15; Gaps 3;

QY 10 VRQERAVRTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTKGAMVFHFPASKEELARGVL 69
DB 27 LKQQRAIRTRGILNRAAFAFATDGPQVTIKDIADGAEMTKGAVYFHPNKEALAVL 86
QY 70 ABQTLHVAVPESGSKAQE-----LVDTLMLVAHGMHLDPILRAGTRLALDQAVDF 120
DB 87 EE--FYRMQEAANGALEHGDPSTPTTVVDMRLARAFHEDVFIHAGARLQIERPYIKA 144
QY 121 SDANFGEWGDIQAOLLAQAQERGEVLPVHNPKKTGDFIVGCTGLQAVSRVTSRDQLG 180
DB 145 ELFPVYVGTGLKVLTELLDQCRTAGNLPKSTKPEALARALGSAVFGAQHISWVINDREDIV 204
QY 181 HRISVMNHNHVLPSIVP 196
DB 205 ERV-----EEIIDA FVP 216

```

```

RESULT 15
US-10-156-761-11240
; Sequence 11240, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11240
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11240

```

```

Query Match      17.8%; Score 207.5; DB 14; Length 287;
Best Local Similarity 24.6%; Pred. No. 5.4e-13;
Matches 70; Conservative 36; Mismatches 70; Indels 109; Gaps 8;

QY 10 VRQERAVRTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTKGAMVFHFPASKEELA----- 65
DB 5 VRQERAIRTNALIESAALFARHGVEVYVLSLTSISARAGVNGALHFFHPKKAALAEAVG 64
QY 66 -----RGVLAEQ-----TLHVAVP 79

```

Search completed: March 6, 2005, 22:37:43
Job time : 130 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 22:12:12 ; Search time 41 Seconds
(without alignments)
544.446 Million cell updates/sec

Title: US-10-049-710A-2
Perfect score: 1164
Sequence: 1 MAVRHERVAVRQERAVRTRQ.....VAAABAAAEAAEAAEAAASDE 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description
Result No.	Score	Query Match %	Length	ID	
1	1164	100.0	232	2 A57507	virginiae butanol
2	363.5	31.2	226	2 T44586	gamma-butyrolacton
3	267	22.9	215	2 T42057	cprA protein - Str
4	257	22.1	215	2 T35863	A-factor receptor
5	220.5	18.9	213	2 T44588	probable transcript
6	169.5	14.6	201	2 AG2376	transcription regu
7	138	11.9	196	2 F90919	repressor for uid
8	138	11.9	196	2 C85768	glucuronide repres
9	138	11.9	196	2 D64918	transcription regu
10	135.5	11.6	203	2 C83773	transcription regu
11	135	11.6	189	2 H96503	probable transcript
12	135	11.6	213	2 H98332	probable transcript
13	134	11.5	197	2 B83485	transcription regu
14	134	11.5	207	2 AC2950	transcription regu
15	133	11.4	190	2 S38906	hypothetical prote
16	132.5	11.4	225	2 C70500	probable transcript
17	132.5	11.4	546	2 G84076	NADH dehydrogenase
18	130.5	11.2	209	2 T34970	probable transcript
19	128.5	11.0	265	2 A87345	transcription regu
20	128	11.0	198	2 E69779	transcription regu
21	127.5	11.0	208	2 E70947	probable helix-tur
22	127.5	11.0	210	2 D87172	transcription regu
23	126.5	10.9	182	2 AD3214	transcription regu
24	126.5	10.9	247	2 AD3649	transcription regu
25	126	10.8	189	2 D70848	probable transcript
26	126	10.8	215	2 A87340	transcription regu
27	124.5	10.7	202	2 B87340	transcription regu
28	124.5	10.7	220	2 D85991	probable transcript
29	124.5	10.7	220	2 B65119	potential acrf/en

30	124.5	10.7	220	2 H91145	probable transcript
31	123.5	10.6	179	2 B70391	transcription regu
32	123.5	10.6	216	2 C70649	hypothetical prote
33	122	10.5	229	2 F75482	transcription regu
34	122	10.5	295	2 AE1398	transcription regu
35	122	10.4	192	2 C70487	transcription regu
36	121	10.4	199	2 D98212	probable tetr-fam
37	121	10.4	207	2 T36937	probable transcript
38	120.5	10.4	202	2 AH1382	transcription regu
39	120.5	10.4	203	2 T36403	probable tetr-fam
40	120.5	10.4	209	2 AI3455	transcription regu
41	119.5	10.3	227	2 S27686	transcription regu
42	119	10.2	195	2 AE3074	transcription regu
43	119	10.2	230	2 A95886	probable transcript
44	118.5	10.2	203	2 S74474	hypothetical prote
45	118	10.1	280	2 C70617	probable transcript

ALIGNMENTS

RESULT 1

A57507
virginiae butanolide-binding protein - Streptomyces virginiae
C;Species: Streptomyces virginiae
C;Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 09-Jul-2004
C;Accession: A57507
R;Okamoto, S.; Nakamura, K.; Nihira, T.; Yamada, Y.
J. Biol. Chem. 270, 12319-12326, 1995
A;Title: Virginiae butanolide binding protein from Streptomyces virginiae. Evidence tha
A;Reference number: A57507; MUID:95263588; PMID:7744885
A;Accession: A57507
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-232 <OKA>
A;Cross-references: UNIPROT:Q60011; GB:D32251; NID:G961502; PIDN:BAA06981.1; PID:d10075
C;Genetics:
A;Gene: bara

Query Match 100.0%; Score 1164; DB 2; Length 232; Best Local Similarity 100.0%; Pred. No. 4.2e-88; Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MAVRHERVAVRQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTKGAMYHFAS	60	
Db	1	MAVRHERVAVRQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTKGAMYHFAS	60	
QY	61	KSELARGVLAEOQLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGA	120	
Db	61	KSELARGVLAEOQLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGA	120	
QY	121	SDANPFGEGWDICQAQLLAEQAERGEVLPVHNPVKKTGDFIVGCTGTGLQAVSRVTS	180	
Db	121	SDANPFGEGWDICQAQLLAEQAERGEVLPVHNPVKKTGDFIVGCTGTGLQAVSRVTS	180	
QY	181	HRISVMNHNVLPSIVPASMLTWIETGEERIGKVAIAAAEAAEAAEAAEAAEASDE	232	
Db	181	HRISVMNHNVLPSIVPASMLTWIETGEERIGKVAIAAAEAAEAAEAAEAAEASDE	232	

RESULT 2

T44586
gamma-butyrolactone receptor protein tyLP [imported] - Streptomyces fradiae
C;Species: Streptomyces fradiae
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T44586
R;Bate, N.; Butler, A.R.; Gandechea, A.R.; Cundliffe, E.
Chem. Biol. 6, 617-624, 1999
A;Title: Multiple regulatory genes in the tylosin-biosynthetic cluster of Streptomyces
A;Reference number: Z22801; MUID:99398833; PMID:10467127
A;Accession: T44586
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

QY	12	QERAVRTRQAI	VRAAASVFDYGFEEAATVABEILSRASVTKGAMYFHFAKSEELARG-VLA	70
Db	6	QTEAQPTRTRIL	NAAREIFSENGFHSASMKAIKCAISPGTLYHHFISKEALIQAILQ	65
QY	71	EQLHVA---	VPESG-----SKAQELVDLT-----MLVAHGMHPILPILRAGTRL	111
Db	66	DQERALARFR	PEPIEGIHFDVYMWVESIVSLTHFAFGQRALVVEIIMAEGM-RNPQVAAMLK-	123
QY	112	ALDQGA	VDFSDANPFGFEGWDICALLAEAOERGBVLPHVNPCKTG- ---DFIVGCTGLQ	167
Db	124	-----	NKHMITEFVAQRMRDAQKGISPDINTAMTSRLLLDLTGYGLVADIE	171
QY	168	A--VSRVTS	DRQDL 179	
Db	172	AEDLAREAS	FAQGL 185	

RESULT 8
 C85768
 repressor for uid operon [imported] - Escherichia coli (strain O157:H7, substrain
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: C85768
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.;
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.;
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: C85768
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-196 <STO>
 A:Cross-references: UNIPROT:Q59431; GB:AE005174; NID:G12515604; PIDN:AA656607.1
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: uidR

Query Match	11.9%	Score 138;	DB 2;	Length 196;
Best Local Similarity	26.3%	Pred. No. 0.00037;		
Matches	51;	Conservative	31;	Mismatches 72;
				Indels 40;
				Gaps 8

QY	12	QERAVRTRQAI	VRAAASVFDYGFEEAATVABEILSRASVTKGAMYFHFAKSEELARG-VLA	70
Db	6	QTEAQPTRTRIL	NAAREIFSENGFHSASMKAIKCAISPGTLYHHFISKEALIQAILQ	65
QY	71	EQLHVA---	VPESG-----SKAQELVDLT-----MLVAHGMHPILPILRAGTRL	111
Db	66	DQERALARFR	PEPIEGIHFDVYMWVESIVSLTHFAFGQRALVVEIIMAEGM-RNPQVAAMLK-	123
QY	112	ALDQGA	VDFSDANPFGFEGWDICALLAEAOERGBVLPHVNPCKTG- ---DFIVGCTGLQ	167
Db	124	-----	NKHMITEFVAQRMRDAQKGISPDINTAMTSRLLLDLTGYGLVADIE	171
QY	168	A--VSRVTS	DRQDL 179	
Db	172	AEDLAREAS	FAQGL 185	

RESULT 9
 D64918
 glucuronide repressor gusR - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C:Accession: D64918
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: D64918
 A:Status: nucleic acid
 A:Molecule type: DNA

A:Residues: 1-196 <BLAT>
A:Cross-references: UNIPROT:Q59431; GB:AE000257; GB:U00096; NID:g1787898; PIDN:AACT74690.
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: gusR; uidR
C:Function:
A:Description: repressor of uidRABC (= gusRABC) operon
C:Keywords: transcription regulation

Query Match 11.9%; Score 138; DB 2; Length 196;
Best Local Similarity 26.3%; Pred. No. 0.00037;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;

QY 12 QBRVTRQAIIVRAAASVDEYGFEEAATVAEILSRASVTGKMGYFHFASKEELARG-VIA 70
DB 6 QTEAQTRILNAAREIPSENGFHSASMKACKSCAISPGTLYHHFISKEALIQAILQ 65

QY 71 EQLHVA-----VPSG-----SKAQLVDLT-----MLVAHGMLHDPILRAGTRL 111
DB 66 DOERALARPREPIEGTHFVDYVNESIVSLTHEAFGORALVVEIMAEGM-RNPQVAAMLK- 123

QY 112 ALDQGAVDPSDANPFGEWGIDICQALLAEQAERGEVLPHVNPCKTG-----DFIVGCTGLQ 167
DB 124 -----NKMWTTEFVAQRMDRAQQGSEISPDINTANTSRLLDLTYGVLAIE 171

QY 168 A--VSRVTSRDQDL 179
DB 172 AEDLAREASFAOGL 185

RESULT 10
C83773
Transcription regulator (Tetr/AcrR family) BH0987 [imported] - Bacillus halodurans (strain C83773)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C83773
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maeui, N.; Fujii, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C83773
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <STO>
A:Cross-references: UNIPROT:Q9KE70; GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BAB047
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0987

Query Match 11.6%; Score 135.5; DB 2; Length 203;
Best Local Similarity 28.5%; Pred. No. 0.00062;
Matches 51; Conservative 22; Mismatches 55; Indels 51; Gaps 7;

QY 7 RVAVROER-AVTRQAIIVRAAASVDEYGFEEAATVAEILSRASVTGKMGYFHFASKEE--- 63
DB 3 RTKEQNERMSMARETKIYKAALSFLFAKGFALTITKIDISREAHISTGLVYRHFHSKEELF 62

QY 64 --LARGVLAEQTLHVAVPESGSKAQLVD-----LTMVAHGMLHDP 103
DB 63 GLRIEKAIGEMTSVTKFLETDTSPELVSEFVTKMIAQSSSEVSHYFLLVARSLLEDE 122

QY 104 IL-RAG-----TRIALDQGAVDPSDANPFGEWGIDICQALLAEQAERGEVLPHVNPCKTG 157
DB 123 VLPKIGDLKKTDIALFPQ-----MAELIARGQRQGEF-----KEGD 158

RESULT 11
H96903
Transcription regulator Tetr/AcrR family [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: H96903
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1C
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: H96903
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-189 <KUR>
A:Cross-references: UNIPROT:Q97N07; GB:AE001437; PIDN:AAK78019.1; PID:g15022852; GSPDB:C
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0032

Query Match 11.6%; Score 135; DB 2; Length 189;
Best Local Similarity 26.7%; Pred. No. 0.00062;
Matches 46; Conservative 31; Mismatches 53; Indels 42; Gaps 7;

QY 17 RTRQAIIVRAAASVDEYGFEEAATVAEILSRASVTGKMGYFHFASKEELARGVLAQTLHV 76
DB 3 KTKNAIFKSALMIFSKNGYDGTWDEIASNAKVAKTLYYHFKSKEEIFKYVISE-GMNV 61

QY 77 AVPSGSKAQLVDLTMLVAHGMLHDP1--LRAGTRLALDQGAV---DFSDANPFGEWG 130
DB 62 IRENEQEA-----GKESNPYNKLAICRFQI--GMIFNRDFFPKVLSQLWG 107

QY 131 -----DICAQLLAQAERGEVLPHVNPCKTGDFIVGCTGL 166
DB 108 QDSRQLELRKVTKDYIEIERYLKDAKEGYI-----KKGETEFMSYTFPGV 154

RESULT 12
H98332
Probable transcription regulator PA3574 [imported] - Agrobacterium tumefaciens (strain C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: H98332
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: H98332
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <KUR>
A:Cross-references: UNIPROT:Q8UB16; GB:AE007870; PIDN:AAK90186.1; PID:g15160193; GSPDB:C
C:Genetics:
A:Gene: AGR_L 3219
A:Map position: linear chromosome
C:Superfamily: probable transcription repressor mtrr

Query Match 11.6%; Score 135; DB 2; Length 213;
Best Local Similarity 31.0%; Pred. No. 0.00072;
Matches 44; Conservative 28; Mismatches 38; Indels 32; Gaps 6;

QY 8 VAVROER-AVTRQAIIVRAAASVDEYGFEEAATVAEILSRASVTGKMGYFHFASKEELA 65
DB 5 ISMRRTKAAEETQAIILAAAEVRFKKGAVNSLDEVAAGVTRGAIYWHFSSKADLF 64

QY 66 RGVLAEQTLHVAVPESGSKAQLVDLTMLVAHGMLHDPILRAGTRLALDQGAVDPSDANP 125
DB 65 IGLYES-----VALPES-----DLID-----FGDPLKGTALLAKIEAT----- 99

QY 126 FGEWGIDICQALLAEQAERGEVL 147
DB 100 -CKW----LALLAEDEBQRORIM 116

RESULT 13
B83485
Probable transcription regulator PA1290 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

This Page Blank (photo)

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	1164	100.0	232	2	Q60011	streptomyc
2	495	42.5	221	2	O24741	streptomyc
3	476	40.9	215	2	O86852	streptomyc
4	470.5	40.4	276	1	AFPP STRGR	streptomyc
5	440	37.8	232	2	O84E47	streptomyc
6	439.5	37.8	234	2	O82H41	streptomyc
7	434.5	37.3	222	2	O83X22	streptomyc
8	428	36.8	237	2	Q6VMG9	streptomyc
9	415	35.7	228	2	O8VVV1	streptomyc
10	403.5	34.7	220	2	O9RPK9	streptomyc
11	402	34.5	208	2	O84H96	streptomyc
12	401	34.5	220	2	Q6X4A0	streptomyc
13	390	33.5	221	2	O6REF8	rhodococc
14	374	32.1	216	2	O6L8I2	kitasatosp
15	363.5	31.2	227	2	Q9XC77	streptomyc
16	338.5	29.1	224	2	Q9ZHP8	streptomyc
17	267	22.9	215	2	O65121	streptomyc
18	257	22.1	215	2	O66122	streptomyc
19	250.5	21.5	230	2	Q6VMH2	streptomyc
20	248.5	21.3	218	2	Q9RSZ2	streptomyc
21	248.5	21.3	218	2	O82H44	streptomyc
22	243	20.9	214	2	Q9JN89	streptomyc
23	236	20.3	228	2	O83X25	streptomyc
24	224	19.2	232	2	O56153	streptomyc
25	224	19.2	225	2	O6REN7	rhodococc
26	220.5	18.9	213	2	O9XCC5	streptomyc
27	217.5	18.7	216	2	O24739	streptomyc
28	215.5	18.5	218	2	O82KU3	streptomyc
29	214	18.4	213	2	O83X30	streptomyc
30	211.5	18.2	224	2	Q93S03	streptomyc
31	210.5	18.1	240	2	O741W9	mycobacteri

```

RESULT 2
O24741
ID O24741 PRELIMINARY; PRT; 221 AA.
AC O24741;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Para.
GN Name=fara;
OS Streptomyces sp.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1931;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRI-5;
RX MEDLINE=97405912; PubMed=9260956;
RA Waki M., Nihira T., Yamada Y.;
RT "Cloning and characterization of the gene (fara) encoding the receptor
RT for an extracellular regulatory factor (IM-2) from Streptomyces sp.
RT strain FRI-5.";
RL J. Bacteriol. 179:5131-5137(1997).
CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
DR EMBL; AB001683; BA021859.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR.1; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 221 AA; 24282 MW; E928FC54AE23P9D7 CRG64;

Query Match 42.5%; Score 495; DB 2; Length 221;
Best Local Similarity 46.8%; Pred. No. 3.4e-32;
Matches 101; Conservative 40; Mismatches 75; Indels 0; Gaps 0;

QY 12 QRAVTRQAIIVRAAASVDEYGFEEATVAEILSRASVTKGMYFHFASKELARGVLA 71
Db 4 QVRAIRTRQAILLSAAARVDEYGFEEATVAEILSRASVTKGMYFHFASKELARGVLA 63

QY 72 QTLHVAVPSSGKAQELVDLTMLVAHGMHDPILRAGTRIALDQGVDFSDANPFGEMGD 131
Db 64 QNEDLLPRLPAKLQEVDAVNLHTRLTNPWVRAGVRLSLDVNAGGLDRSAPFNWVD 123

QY 132 ICAQLLAEAQERGEVLPHVNPVKTKGDFIVGCTGLQAVSRVTSRDQDLGHRISVMNVHL 191
Db 124 KETDLEKAQAQCELLPHVVPAETADVTGAYGVQSMQALTEHQDLGQRVNALRLHLM 183

QY 192 PSIVPASMLTWIETGEERIGKVAARAAAEAAEASE 227
Db 184 PSIAQPSVLASLHGESAEVLEARQLARQADE 219

RESULT 3
O86852
ID O86852 PRELIMINARY; PRT; 215 AA.
AC O86852;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Gamma-butyrolactone binding protein.
GN Name=schr; ORFNames=SCAH10.30c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Eubacteria, and M145;

Takano E., Chakraborty R., Nihira T., Yamada Y., Bibb M.;
Submitted (AUG-1998) to the EMBL/GenBank/DBSJ databases.
RA
RL
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
DR EMBL; AJ007731; CAA07628.1; -.
DR EMBL; AL939127; CAB60184.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR.1; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 215 AA; 23861 MW; C5D30C679408ABB9 CRC64;

Query Match 40.9%; Score 476; DB 2; Length 215;
Best Local Similarity 47.3%; Pred. No. 1.1e-30;
Matches 97; Conservative 36; Mismatches 72; Indels 0; Gaps 0;

QY 11 RQERAVTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTKGMYFHFASKELARGVLA 70
Db 3 KQDAIRTRQITLDAQAQVFEKQGYQAATITELKLVAGVTKGALYFHFQSKELALGVDF 62

QY 71 EQLHVAVPSSGKAQELVDLTMLVAHGMHDPILRAGTRIALDQGVDFSDANPFGEMG 130
Db 63 AQEPVQAVPEQPLRIQELIDMGMLFCHRLRTNVVARAGVRLSMDQQAHLDRRGPFRRWH 122

QY 131 DTCAQLLAEAQERGEVLPHVNPVKTKGDFIVGCTGLQAVSRVTSRDQDLGHRISVMNVH 190
Db 123 ETLLKLLNQAKENGELLPHVVTDSADLYVGTGTFAGIQVVSQVSDYQDLGHRVALLQXHI 182

QY 191 LPSIVPASMLTWIETGEERIGKVA 215
Db 183 LPAIAVPSVLAALDLSEERGARLAA 207

RESULT 4
AFRP_STRGR
ID AFRP_STRGR STANDARD; PRT; 276 AA.
AC Q9ZNF78; Q54189;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE A-factor receptor protein (A-factor binding protein).
DE Name=arpa;
GN Streptomyces griseus.
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.; SEQUENCE OF 6-21; 44-53 AND 148-167, SUBUNIT.
RP A-FACTOR-BINDING, AND VARIANTS.
RC STRAIN=IFO 13350;
RX MEDLINE=96042082; PubMed=7592371;
RA Onaka H., Ando N., Nihira T., Yamada Y., Beppu T., Horinouchi S.;
RT "Cloning and characterization of the A-factor receptor gene from
```



```

OX NCBI_TaxID=1901;
RN SEQUENCE FROM N.A.
RX PubMed=15257430;
RA Kim H.S., Lee Y.J., Lee C.K., Choi S.U., Yeo S.H., Hwang Y.I.,
RA Yu T.S., Kinoshita H., Nihira T.;
RT "Cloning and characterization of a gene encoding the gamma-
RT butyrolactone autoregulator receptor from Streptomyces clavuligerus.";
CC Arch. Microbiol. 182:44-50(2004).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL: AB106894; BAC66444.1; -.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR009050; Globin like.
DR InterPro: IPR001647; HTH Tetr.
DR Pfam: PF00440; Tetr_N; 1.
DR PRINTS: PR00455; HTHETR.
DR PROSITE: PS01081; HTH TETR 1; 1.
KW Complete proteome, DNA-binding; Receptor; Transcription;
SQ SEQUENCE 232 AA; 25423 MW; 4EAB640486C7B27B CRC64;

Query Match 37.8%; Score 440; DB 2; Length 232;
Best Local Similarity 45.9%; Pred. No. 1e-27;
Matches 94; Conservative 40; Mismatches 65; Indels 6; Gaps 4;

QY 11 ROEAVRTQALVRAAASVFDEYGEAATVAETLSRASVTGKAMYPHFASKELARGVLA 70
DB 3 ROEALRTKRLVVAASVFDQVGEAATISDLQSGLTGKALYPHFGSKELAQAVLA 62
QY 71 EQLTHV-APESGSKAQELVDLTMLVAH---GMLHDPILRAGTRALALDQG-AVDFSDAN- 124
DB 63 EQVAALPPVPQGLKLEAIDEGMLLAHLTGEHGDPIIQSGIRLTVDQSSKNDLDRV 122
QY 125 PGEWGDICQAQLAAEQAERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLGHRIS 184
DB 123 PMQGWTDHSLGLFSEARKNGEVLPHADLDSVSRFLFAGCYTGQVLSRIMTRDQDLAERMS 182
QY 185 VMNHNVLPSIVPASMLTWIETGEER 209
DB 183 DMIRNLMPALAVPSVLRLDFTSDT 207

RESULT 6
Q82H41 PRELIMINARY; PRT; 234 AA.
AC Q82H41;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative gamma-butyrolactone receptor protein.
GN Name=avar; OrderedLocusNames=SAV3705;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

SAKAKI Y., HATTORI M., OMURA S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL: AP005036; BAC71417.1; -.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR009050; Globin like.
DR InterPro: IPR009057; Homeodomain like.
DR InterPro: IPR001647; HTH Tetr.
DR Pfam: PF00440; Tetr_N; 1.
DR PRINTS: PR00455; HTHETR.
DR PROSITE: PS01081; HTH TETR 1; 1.
KW Complete proteome, DNA-binding; Receptor; Transcription;
SQ SEQUENCE 234 AA; 25833 MW; 7773764536A80CF2 CRC64;

Query Match 37.8%; Score 439.5; DB 2; Length 234;
Best Local Similarity 42.6%; Pred. No. 1.2e-27;
Matches 95; Conservative 44; Mismatches 77; Indels 7; Gaps 4;

QY 11 ROEAVRTQALVRAAASVFDEYGEAATVAETLSRASVTGKAMYPHFASKELARGVLA 70
DB 3 ROEALRTKRLVVAASVFDQVGEAATISDLKRSRGVTKGALYFHTSKQELAQAVLA 62
QY 71 EQLTHV-APESGSKAQELVDLTMLVAHGMHLH--DPILRAGTRALALDQGV--DFSDAN 124
DB 63 EQVASLPRVPEQLKQSLDGLALHLRBTGDPVQGSVRVTVDQGSPRDLHNRV 122
QY 125 PGEWGDICQAQLAAEQAERGEVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDQDLGHRIS 184
DB 123 PMQWTEHTQSIFEEARAKGEILPHADVEALAKLVGAGTGVQVLSRIMTRGADLAERVA 182
QY 185 VMNHNVLPSIVPASMLTWIETGEERIGKV-AAAAEAAEAAEAS 226
DB 183 DLYRLMPSFAMPGLIVRLDFSPERSGRVYEAAMQRESAAS 225

RESULT 7
Q83X22 PRELIMINARY; PRT; 222 AA.
AC Q83X22;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative tetr family receptor protein.
OS Streptomyces rochei (Streptomyces parvullus).
OG Plasmid pSLA2-L.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=22676866; PubMed=12791134;
RA Mochizuki S., Hiratsu K., Suwa M., Ishii T., Sugino F., Yamada K.,
RA Kinashi H.;
RT "The large linear plasmid pSLA2-L of Streptomyces rochei has an
RT unusually condensed gene organization for secondary metabolism.";
RL Mol. Microbiol. 48:1501-1510(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=20408175; PubMed=10954087;
RA Hiratsu K., Mochizuki S., Kinashi H.;
RT "Cloning and analysis of the replication origin and the telomeres of
RT the large linear plasmid pSLA2-L in Streptomyces rochei.";
RL Mol. Gen. Genet. 263:1015-1021(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;

```

RX MEDLINE=20331737; PubMed=10767533; DOI=10.1016/S0378-1119(00)00060-3;
RA Suwa M., Sugino H., Sasaka A., Mori E., Fujii S., Shinkawa H.,
RA Nimi O., Kinashi H.,
RT "Identification of two polyketide synthase gene clusters on the linear
RT plasmid pSLA2-L in Streptomyces rochei";
RL Gene 246:123-131(2000).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=7434AN4;
RC MEDLINE=9053144; PubMed=9836424;
RA Kinashi H., Fujii S., Hatanai A., Kurokawa T., Shinkawa H.;
RT "Physical mapping of the linear plasmid pSLA2-L and localization of
RT the eryAI and actI homologs";
RL Biotechnol. Biochem. 62:1892-1897(1998).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AB088224; BAC76540.1; -.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001647; HTH Tetr.
DR PROSITE; PS01081; HTH Tetr.
DR Pfam; PF00440; Tetr.N.1.
DR PRINTS; PR00455; HTH Tetr.
DR PROSITE; PS01081; HTH TETR_1; 1.
DR DNA-binding; Transcription; Transcription regulation.
KW Transcription regulation.
SQ SEQUENCE 222 AA; 23805 MW; EBE16185AFB85F09 CRC64;

Query Match 37.3%; Score 434.5; DB 2; Length 222;
Best Local Similarity 42.7%; Pred. No. 2.8e-27;
Matches 94; Conservative 45; Mismatches 78; Indels 3; Gaps 3;

QY 11 ROERAVRTRQAIIVRAAASVFDEYGFPAATVAEILSRASVTGKAMYPHFASKELARGVLA 70
Db 3 QGERAIRTRRAVLEAAATVFAEYGAATVADILKVAGLTKGALYFHPFSKEALARGILE 62

QY 71 EQTLHVAVPESGSKAQELVDLTMLVAHGMHDPILRAGTRRLALDQGVDFSDANPFGW 130
Db 63 AQVPOQLVPQQ-LKIQEWDAGMTLAHQLPDRPVRVAGARLSAETHGTSE-QHGSAPPTWI 120

QY 131 DICAOALLAAOERGEVLPVHNPKTGDFIVGCTGLQAVSRVTSDDQDLGHRISVVMNHV 190
Db 121 AFSASLLEQKNGEVGLHIEPAETACVIGSGFHGLQLLSQTNWADIEQASALFRHV 180

QY 191 LPSIVPASMLTWTETGEERIGKVAASAAEAAEASEAAS 230
Db 181 LPAAVPSVLVRLDTAPDRGARVVAELEAM-AAQPDGLAS 219

RESULT 8
Q6VMG9
ID Q6VMG9 PRELIMINARY; PRT; 237 AA.
AC Q6VMG9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative gamma-butyrolactone binding protein.
GN Name=alpZ;
OS Streptomyces ambofaciens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1889;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23877;
RX PubMed=14742212;
RA Fang X., Aigle B., Girardet J.M., Mangenot S., Pernodet J.L.,
RA Decaris B., Leblond P.;
RT "Functional anacycline-like antibiotic gene cluster in the terminal
RT inverted repeats of the Streptomyces ambofaciens linear chromosome.";
RL Antimicrob. Agents Chemother. 48:575-588(2004).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AY338477; AAR30170.1; -.

DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001647; HTH Tetr.
DR InterPro; IPR001647; HTH Tetr.
DR Pfam; PF00440; Tetr.N.1.
DR PRINTS; PR00455; HTH Tetr.
DR PROSITE; PS01081; HTH TETR_1; 1.
DR DNA-binding; Transcription; Transcription regulation.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 237 AA; 25701 MW; F649FA2A64AE5834 CRC64;

Query Match 36.8%; Score 428; DB 2; Length 237;
Best Local Similarity 40.5%; Pred. No. 1e-26;
Matches 90; Conservative 47; Mismatches 85; Indels 0; Gaps 0;

QY 11 ROERAVRTRQAIIVRAAASVFDEYGFPAATVAEILSRASVTGKAMYPHFASKELARGVLA 70
Db 3 QGERAIRTRRAVLEAAAEVFDGSHGYEATIGELIRAGVTGKALYFHPFSKQALAEVLE 62

QY 71 EQTLHVAVPESGSKAQELVDLTMLVAHGMHDPILRAGTRRLALDQGVDFSDANPFGW 130
Db 63 QQFSVIRVPPGFCRLQEFVDTGLIVAYRMRDPLVSAVARLSLEQELRAEYSGSAIRQWI 122

QY 131 DICAOALLAAOERGEVLPVHNPKTGDFIVGCTGLQAVSRVTSDDQDLGHRISVVMNHV 190
Db 123 GASEVLGAAKEQGLHLPVPAESANLPSAAWTGTQLYSQILLGREDLEERVVALFRHL 182

QY 191 LPSIVPASMLTWTETGEERIGKVAASAAEAAEASEAASDE 232
Db 183 LPSIAPVAVLSLEITERRAARLGSASDALLGPGEQADSPQ 224

RESULT 9
Q8VVP1
ID Q8VVP1 PRELIMINARY; PRT; 228 AA.
AC Q8VVP1
DT 01-VAR-2002 (TrEMBLrel. 20, Created)
DT 01-VAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-VAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Autoregulator receptor protein.
GN Name=spBR;
OS Streptomyces pristinaespiralis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=38300;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21570198; PubMed=11557748; DOI=10.1074/jbc.M101109200;
RA Folcher M., Gaillard H., Nguyen L.T., Nguyen K.T., Lacroix P.,
RA Bamas-Jacques N., Rinkel M., Thompson C.J.;
RT "Pleiotropic functions of a Streptomyces pristinaespiralis
RT autoregulator receptor in development, antibiotic biosynthesis, and
RT expression of a superoxide dismutase";
RL J. Biol. Chem. 276:44297-44306(2001).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AY026762; AAK07686.1; -.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF00440; Tetr.N.1.
DR PRINTS; PR00455; HTH Tetr.
DR PROSITE; PS01081; HTH TETR_1; UNKNOWN_1.
DR DNA-binding; Receptor; Transcription; Transcription regulation.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 228 AA; 25012 MW; ICC9EF160C756CB5 CRC64;

Query Match 35.7%; Score 415; DB 2; Length 228;
Best Local Similarity 43.9%; Pred. No. 1.1e-25;
Matches 98; Conservative 39; Mismatches 76; Indels 10; Gaps 6;

QY 11 ROERAVRTRQAIIVRAAASVFDEYGFPAATVAEILSRASVTGKAMYPHFASKELARGVLA 70
Db 3 QGERAVRTRRAVLEAAAEVFDGSHGYEATIGELIRAGVTGKALYFHPFSKQALAEVLE 62

QY 71 EQTLHVA--VPESGSKAQELVDLTMLVAHGMH---DPILRAGTRRLALDQGVDFSDAN 124


```
Db 63 EQ-VHIALPOLPEGLMLQTAVDRAILLALHLLRRDTGDPVIRGSLVLTVEQALRDGLDR 121
QY 125 -PFGWGDIACAOLLAERGERVLPVHNPKTKGDFIVGCTGLQAVSRVTSRDOLGHR 183
Db 122 VPMQAMEGTQDLFOQAAGSILPHVDLIVGAAKTFVGAGTGVQLSNIMTGRQDMTERV 181
QY 184 SVMNHVLPISVPASMLTWIETGEERIGKVAIAAAAEAAAEAS 226
Db 182 ADLYRFLMTAIVPGVLVLDSPGR--GVLAEEAVERDDAA 222

RESULT 10
Q9RPK9 ID Q9RPK9 PRELIMINARY; PRT; 220 AA.
AC Q9RPK9
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TarA.
OS Streptomyces tendae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1932;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC31160;
RA Engel P. Scharfenstein L.L.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
EMBL; AF156163; AAF06961.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; UNKNOWN 1.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 220 AA; 24010 MW; DCEABE24C6581D02 CRC64;
```

```
Query Match 34.7%; Score 403.5; DB 2; Length 220;
Best Local Similarity 44.0%; Pred. No. 9e-25;
Matches 92; Conservative 41; Mismatches 71; Indels 5; Gaps 4;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGEPAATVAETLSRASVTKGAMYPHFASKKEELARGVLA 70
Db 3 QODRAVTRRAVIRAAAVFAERGAAATISEILKRAGVTGKALYFHFDSKAALAQGV 62

QY 71 EQ-TLHVAVPESGSKAQELVDLTMLVAHGMLHDPILRAGTRIALDO-GAVDFSDANPFGE 128
Db 63 EQLTPEYHLPRE-LKLQEWVDAGMTLARRLPREFLLAGVRISADRPGRGVLSAWP--A 119

QY 129 WGDICQALLAEARQERGVLPVHNPKTKGDFIVGCTGLQAVSRVTSRDOLGHRISVMWN 188
Db 120 WSLRSHVLTAKRQERGVLPVHNPKTKGDFIVGCTGLQAVSRVTSRDOLGHRISVMWN 179

QY 189 HVLPSVTPASMLTWIETGEERIGKVAIAAA 217
Db 180 HLLGAIAAPVLTLDTPADRGARVIAEA 208
```

```
RESULT 11
Q84H96 ID Q84H96 PRELIMINARY; PRT; 208 AA.
AC Q84H96
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Butyrolactone receptor.
OS Streptomyces carzinostaticus subsp. neocarzinostaticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
```

```
OX NCBI_TaxID=167636;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15944;
RA Liu W., Nonaka K., Nie L., Bae J., Zazopoulos E., Parnet C.M.,
RA Shen B.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
EMBL; AV117439; AAM78022.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; UNKNOWN 1.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 208 AA; 23076 MW; F328B30B25E7507B CRC64;
```

```
Query Match 34.5%; Score 402; DB 2; Length 208;
Best Local Similarity 42.5%; Pred. No. 1.1e-24;
Matches 79; Conservative 37; Mismatches 70; Indels 0; Gaps 0;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGEPAATVAETLSRASVTKGAMYPHFASKKEELARGVLA 70
Db 5 QODRAVTRRAVIRAAAVFAERGAAATISEILKRAGVTGKALYFHFDSKAALAQGV 64

QY 71 EQTLHVAVPESGSKAQELVDLTMLVAHGMLHDPILRAGTRIALDOGAVDFSDANPFGE 130
Db 65 SQOTFIDFPSEGAGLQVRVIDLTFLHAHELQTNPLIRASIRLAVEQGEFGVRDQDTAYQDW 124

QY 131 DTCQAQLLAERQERGVLPVHNPKTKGDFIVGCTGLQAVSRVTSRDOLGHRISVMWNHV 190
Db 125 ALVETYLLEARRQERGVLPVHNPKTKGDFIVGCTGLQAVSRVTSRDOLGHRISVMWNHV 184

QY 191 LPSIVP 196
Db 185 LPGLTP 190
```

```
RESULT 12
Q6X4A0 ID Q6X4A0 PRELIMINARY; PRT; 220 AA.
AC Q6X4A0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative gamma-butyrolactone-binding protein.
OS Streptomyces ansochromogenes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=115647;
RN 11
RP SEQUENCE FROM N.A.
RA Li W., Tan H.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
EMBL; AY256849; AAP87374.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; 1.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 220 AA; 23920 MW; B76FF63127595602 CRC64;
```

```
Query Match 34.5%; Score 401; DB 2; Length 220;
Best Local Similarity 42.5%; Pred. No. 1.4e-24;
Matches 94; Conservative 40; Mismatches 79; Indels 8; Gaps 4;
```

```
QY 11 RQRAVTRQAIIVRAAASVDFEYGFPAATVAEILSPASVTKGAMYFHPFASKKEELARGVLA 70
Db 3 QODRAVTRRAVRAAASVFAERGAAYATSEILKAGVTKGALYFHPFASKKEELARGVLAQ 62
QY 71 EQ-TLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRALDQAGVDFSDANPFGW 129
Db 63 EQITPEYHLPRE-LKLEWVDAGMALARLPRPFFLLAGVRISADPRGDVL-GSAPWPAW 120
QY 130 GICACALLAQAQRGEVLPHPNPKTGDPIVCGFTGLQAVSRVTSRDQLGHRISVMWNH 189
Db 121 ARLTSHALTEAKRGEVLPHPNPKTGDPIVCGFTGLQAVSRVTSRDQLGHRISVMWNH 189
QY 190 VLPSIVPASMLTWIEGGERIGKVAASAEAAEAAEAAEAAEAAEAAEAAEAAEAAEAAE 230
Db 181 LLAAIAAPPVLTLDTPDRGARV-----IAEARRESEDL 216

RESULT 13
Q6REF8 PRELIMINARY; PRT; 221 AA.
AC Q6REF8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PK3.089;
OS Rhodococcus sp. DK17.
OG Plasmid pDK3.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=186196;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DK17;
RA Choi K.Y., Kim D., Sul W.J., Chae J.-C., Zylstra G.J., Kim Y.M.,
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AY502076; AAR90230.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTEPR.
KW DNA-binding; Hypothetical protein; Plasmid; Transcription;
KW Transcription regulation.
SQ SEQUENCE 221 AA; 24204 MW; B52906C946B94CAE CRC64;

Query Match 33.5%; Score 390; DB 2; Length 221;
Best Local Similarity 40.5%; Pred. No. 1.le-23;
Matches 79; Conservative 39; Mismatches 71; Indels 6; Gaps 1;

QY 10 VROEAVRTRQAIIVRAAASVDFEYGFPAATVAEILSPASVTKGAMYFHPFASKKEELARGV 69
Db 2 VQOQRAVTRQAIIVRAAASVDFEYGFPAATVAEILSPASVTKGAMYFHPFASKKEELARGV 61
QY 70 AEQ-----TLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRALDQAGVDFSDA 123
Db 62 AEQHRISISAVQAIAQAESAIEQIVMLCHEMARQIVQDPPIVRAGIRITLELSADDRGPA 121
QY 124 NPFGWGDICACALLAQAQRGEVLPHPNPKTGDPIVCGFTGLQAVSRVTSRDQLGHRIS 183
Db 122 GPYLDWIAQCQLARAVAEGLVDITDPTFARYVIGATGVTQVTSQVLTTHRADLEQRV 181
QY 184 SVMNHNVLPSIVPAS 198
Db 182 DEMWGFLLPGIMPTT 196

RESULT 14
Q6L812 PRELIMINARY; PRT; 217 AA.
ID Q6L812
```

```
AC Q6L812;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE KsBA (Fragment).
GN Name=kSBA;
OS Kitasatospora setae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Kitasatospora.
OX NCBI_TaxID=2066;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14216; and IFO14216;
RX PubMed=15150228; DOI=10.1128/JB.186.11.3423-3430.2004;
RA Choi S., Lee C., Hwang Y., Kinoshita H., Nihira T.;
RT "Cloning and functional analysis by gene disruption of a gene encoding
a gamma-butyrolactone autoregulator receptor from Kitasatospora
setae.";
RL J. Bacteriol. 186:3423-3430(2004).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AB126048; BAD20233.1; -.
DR EMBL; AB121071; BAD20239.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTEPR.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 217 AA; 23755 MW; 54C7579CCF7C33DA CRC64;

Query Match 32.1%; Score 374; DB 2; Length 217;
Best Local Similarity 35.2%; Pred. No. 2.2e-22;
Matches 77; Conservative 51; Mismatches 81; Indels 10; Gaps 2;

QY 14 RAVTRQAIIVRAAASVDFEYGFPAATVAEILSPASVTKGAMYFHPFASKKEELARGVLA 73
Db 6 RAKTREALIQAAAGVFEQHGNGARLSMITSRSLGTWGAIVFHTSKDDLARAVNAAQA 65
QY 74 LHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRALDQAGVDFSDANPFGWGDIC 133
Db 66 DDLVLPDGERGLQRLDITLYLASELQRLNVLVRAGVRLAEQSGFGRVDRSTPYLQWVDRF 125
QY 134 AQLLAQAQRGEVLPHPNPKTGDPIVCGFTGLQAVSRVTSRDQLGHRISVMNHNVLPS 193
Db 126 RDELGAEEAVGOLLPEVVAVQALLVSSYSTQLLSEIATDRADLPRIVRMWTYLLPG 185
QY 194 IVPASMLTWIEGGERIGKVAASAEAAEAAEAAEAAEAAEAAEAAEAAEAAEAAE 232
Db 186 IA-----TPEART-RLRLEPSVADRDADRGNDQ 214

RESULT 15
Q9XCX7 PRELIMINARY; PRT; 226 AA.
ID Q9XCX7
AC Q9XCX7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gamma-butyrolactone receptor protein TyLP.
GN Name=tyLP;
OS Streptomyces fradiae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1906;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T59235;
RX MEDLINE=993998833; PubMed=10467127; DOI=10.1016/S1074-5521(99)80113-6;
RA Bate N., Butler A.R., Gandeche A.R., Cundliffe E.;
RT "Multiple regulatory genes in the tylosin biosynthetic cluster of
Streptomyces fradiae.";
RL Chem. Biol. 6:617-624(1999).
```

